ELSEVIER

A spatial–temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery

Desheng Liu [a,*], Maggi Kelly [a], Peng Gong [a,b]

[a] Center for Assessment and Monitoring of Forest and Environmental Resources, Department of Environmental Science,
Policy, and Management, University of California at Berkeley, 137 Mulford Hall # 3114, CA 94720-3110, USA
[b] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and
Beijing Normal University, Beijing, 100101, China

Abstract

Sudden Oak Death is a new and virulent disease affecting hardwood forests in coastal California. The spatial–temporal dynamics of oak mortality at the landscape scale are crucial indicators of disease progression. Modeling disease spread requires accurate mapping of the dynamic pattern of oak mortality in time through multi-temporal image analysis. Traditional mapping approaches using per-pixel, single-date image classifications have not generated consistently satisfactory results. Incorporation of spatial–temporal contextual information can improve these results. In this paper, we propose a spatial–temporally explicit algorithm to classify individual images using the spectral and spatial–temporal information derived from multiple co-registered images. This algorithm is initialized by a spectral classification using Support Vector Machines (SVM) for each individual image. Then, a Markov Random Fields (MRF) model accounting for ecological compatibility is used to model the spatial–temporal contextual prior probabilities of images. Finally, an iterative algorithm, Iterative Conditional Mode (ICM), is used to update the classification based on the combination of the initial SVM spectral classifications and MRF spatial–temporal contextual model. The algorithm was applied to two-year (2000, 2001) ADAR (Airborne Data Acquisition and Registration) images, from which three classes (bare, dead, forest) are detected. The results showed that the proposed algorithm achieved significantly better results (Year 2000: Kappa=0.92; Year 2001: Kappa=0.91), compared to traditional pixel-based single-date approaches (Year 2000: Kappa=0.67; Year 2001: Kappa=0.66). The improvement from the contributions of spatial–temporal contextual information indicated the importance of spatial–temporal modeling in multi-temporal remote sensing in general and forest disease modeling in particular.
© 2005 Elsevier Inc. All rights reserved.

Keywords: High spatial resolution; Sudden Oak Death; Multi-temporal imagery; Spatial–temporal classification; Support Vector Machines; Markov Random Fields

1. Introduction

In central coastal California, a newly discovered pathogen *Phytophthora ramorum* has been killing hundreds of thousands of tanoak (*Lithocarpus densiflorus*), coast live oak (*Quercus agrifolia*), and black oak (*Quercus kelloggii*) trees (Rizzo et al., 2002). Infected trees can take between 2 and 20 years to die (McPherson et al., 2005), however in the last years of life infected trees decline rapidly, with in many cases the whole canopy changing color. This change in appearance has lead to the disease complex name "sudden oak death" (SOD) (Rizzo & Garbelotto, 2003). Oaks are a major component of many California hardwood forest ecosystems, and the trees are also important both in urban landscapes and at the urban/rural interface (Garbelotto et al., 2001). Currently, the disease has officially been confirmed in 14 coastal counties of California, and one county in Oregon. It has reached epidemic proportions in selected forest areas from Monterey County to Curry County in southern Oregon (Rizzo & Garbelotto, 2003), and hosts for the disease exist across both states and exist elsewhere nationally. The disease was first reported in 1995 in Marin County, CA and forests there, including the mixed hardwood forests found in China Camp State Park (CCSP) display extensive overstory mortality of coast live oaks and black oaks (Kelly et al., 2004a,b; Rizzo & Garbelotto, 2003).

---

* Corresponding author.
  E-mail address: dsliu@nature.berkeley.edu (D. Liu).

Spatial pattern is one of the most fundamental properties of disease dynamics because it reflects the environmental forces acting on the dispersal and life cycles of a pathogen (Ristaino & Gumpertz, 2000). For this reason, researchers of plant disease epidemics are increasingly using landscape approaches to quantify and model spatial patterns of disease spread in order to understand the basic factors that influence pathogen dispersal and infection processes (Holdenrieder et al., 2004; Wulder et al., 2004). Remote sensing is often used in this context, and while remotely sensed approaches to monitoring individual tree health were developed earlier in temperate conifer forests, examples of such work in other forest systems, including hardwood forests are increasing (Asner et al., 2004; Boyer et al., 1988; Clark et al., 2004a,b; Everitt et al., 1999; Gong et al., 1999; Key et al., 2001; Muchoney & Haack, 1994). While remote sensing may not be able to discern the cause of death of a tree, a remote approach to mapping mortality in a forest facing an epidemic remains important for disease monitoring, epidemiology and pathology. Therefore, it is desirable to develop an automated remote sensing monitoring approach based on multi-temporal high spatial resolution imagery.

Previous efforts in mapping the oak mortality dynamics through classifying multi-temporal high spatial resolution imagery were mainly focused on non-contextual classification algorithms, for example, maximum likelihood classification and ISODATA clustering (Kelly, 2002; Kelly et al., 2004a). The initial results indicated that these algorithms were insufficient to generate consistently satisfactory results because of several complicating factors associated with the vegetation being mapped (Kelly, 2002; Kelly et al., 2004a). First, there exists considerable spectral confusion between dead trees and bare areas, especially for bare areas at forest edges, where grass, other understory vegetation and bare soils are mixed. As a result, some bare areas can be misclassified as dead trees. Second, seasonal differences in phenology have also been problematic when mapping hardwood mortality through time. Specifically, when imagery is collected at the time that deciduous oaks have leafed out with immature leaves, the leaves, while green, still have a depressed NIR reflectance. Thus deciduous oaks (primarily valley oaks) can appear very similar to dead trees in classification routines, resulting in commission errors. Third, as a result of non-contextual classification with high spatial resolution imagery, the classification maps included segmented and isolated pixels. These problems in classification are compounded when comparing classification results from two time periods. Considerable artifacts and spurious changes can be observed due to independent classification errors and registration errors, which make the subsequent landscape level spatial–temporal dynamics analysis unreliable and difficult.

The technical limitations of conventional approaches for mapping SOD dynamics with high spatial resolution imagery can be attributed to their simplified assumptions: spatial independence among individual pixels and temporal independence between images from different dates. With these assumptions, multiple image series from different time periods are separately processed without referring to any useful temporal relation, and individual pixels on each image are independently treated without considering their spatial association (Kelly, 2002; Kelly et al., 2004a). However, on high spatial resolution imagery, spatially neighboring pixels tend to belong to the same class or some compatible classes with an ecological association. For a series of multiple images, temporally neighboring pixels (Refer to Section 2.2) are related to each other through phenology, land use, or other rationales. Therefore, pixels on multi-temporal high spatial resolution imagery show dependencies in both the spatial and temporal domains. Ignoring such spatial–temporal dependencies associated with multi-temporal high spatial resolution imagery does not maximize the information contained in the datasets and may result in an unsatisfactory classification result. This is especially true when spectral information has limited discriminative power where spatial–temporal dependencies can help to remove the spectral confusion. Therefore, it is likely that by modeling spatial–temporal dependencies in the classification process the accuracy of the classified map will be increased due to the incorporation of complimentary contextual information.

The use of either spatial or temporal information in remote sensing imagery has been extensively explored and shown to generate improved results over methods based on spectral data alone (Bruzzone & Smits, 2001; Gong & Xu, 2003; Richards & Jia, 1999). In contrast, the use of both spatial and temporal information is rather limited (Jeon & Landgrebe, 1999; Liu et al., 2005; Melgani & Serpico, 2003; Solberg et al., 1996). For monitoring the dynamics of SOD, a spatial–temporal approach is necessary, considering the high spatial resolution of the data and the multi-temporal nature of the problem. In this paper, we propose a spatial–temporally explicit algorithm based on Markov Random Fields (MRF) and Support Vector Machines (SVM) to classify multi-temporal images simultaneously for SOD dynamics monitoring. We have two objectives: 1) developing a new spatial–temporal classification method for the use of multi-temporal remote sensing data and comparing it with conventional non-contextual approaches, 2) applying the new method in a case study of monitoring the dynamics of SOD and discussing the broader application of this new method. The reminder of the paper is organized as follows. In Section 2, we provide some theoretical background on SVM and MRF, which are the building blocks of our proposed algorithm. In Section 3, we describe the study site and data. In Section 4, we present our spatial–temporally explicit algorithm. In Sections 5 and 6, we show the results and together with some discussion on several important issues. We conclude in Section 7.

2. Theoretical background

2.1. Support Vector Machines

Support Vector Machines (SVM) are kernel-based learning classifiers built on statistical learning theory (Burges, 1998; Cristianini & Shawe-Taylor, 2001; Vapnik, 1998). SVM are trained to find the optimal classification hyperplane minimizing an upper bound of the generalization error. This makes SVM more attractive than other classifiers in a broad range of

applications. For remotely sensed imagery, SVM have been found to be competitive with the best available machine learning algorithms in classifying high-dimensional data sets (Huang et al., 2002).

Considering a binary classification problem, the training samples consist of a data set

$$\{(x_i, y_i) | x_i \in R^d, y_i \in \{-1, 1\}, i = 1, \ldots, n\}$$

where x_i is a d-dimensional vector and y_i is the corresponding class label. A supervised training algorithm aims to find a decision function

$$g(x) = sign(f(x)) \tag{1}$$

from the known training samples based on certain learning rules. For linearly separable cases, a linear SVM is designed to derive an optimal linear discriminant function

$$f(x) = w^T x + b \tag{2}$$

where $w = [w_1, w_2, \ldots, w_d]^T$ is the weight vector and b is a bias term. The linear separating hyperplane is characterized by $w^T x + b = 0$. The weight vector w is optimized so that the margin between linear boundary $w^T x + b = 1$ and $w^T x + b = -1$ is maximized. Mathematically, the primal optimization problem is formulated as:

$$w^* = \arg\min_w \left\{ \frac{1}{2} \|w\|^2 \right\} \text{ subject to } \{y_i(w^T x_i + b) \geq 1, i = 1, 2 \ldots, n\}. \tag{3}$$

Using Lagrange formulation, the optimal discriminant function can be expressed in terms of Lagrangian dual variables α_i:

$$f(x) = \sum_{i \in SV} \alpha_i y_i \langle x_i, x \rangle + b \tag{4}$$

where SV (Support Vector) is the set of training samples with associated dual variables α_i satisfying $\alpha_i > 0$.

For linearly nonseparable cases, slack variables $\xi_i (i = 1, 2, \ldots, n)$ are introduced to penalize the misclassification errors and relax the hard constraints in Eq. (3), then the primal optimization problem becomes

$$w^* = \arg\min_w \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right\}$$

subject to $\{y_i(w^T x_i + b) \geq 1 - \xi_i, i = 1, 2, \ldots, n\}$. \tag{5}

With Lagrange formulation, the same form of optimal discriminant function as in Eq. (4) can be obtained except that the Lagrangian dual variables α_i are enforced with an upper bound C, the regularization parameter. In these cases, the support vectors consist of a subset of training samples with associated dual variables α_i satisfying $0 < \alpha_i < C$.

In cases of nonlinear classification boundaries, the above formulations are generalized to their nonlinear counterparts by transforming the original input space nonlinearly to a higher dimensional feature space where linear methods may be applied. In doing so, a kernel trick is introduced into SVM so that the discriminant function in Eq. (4) becomes:

$$f(x) = \sum_{i \in SV} \alpha_i y_i K(x_i, x) + b \tag{6}$$

By substituting the standard dot product $\langle x_i, x \rangle$ with a kernel function $K(x_i, x)$, SVM implicitly map the original data from the input space R^d to a higher dimensional feature space H via a nonlinear mapping function φ such that

$$K(x_i, x) = \langle \varphi(x_i), \varphi(x) \rangle \tag{7}$$

In this way, SVM find an optimal linear hyperplane in a higher dimensional feature space that is nonlinear in the original input space. The kernel trick avoids the direct evaluation of the dot product in the higher dimensional feature space by computing it via kernel function with data vectors in the input space. A valid kernel function must obey the Mercer's theorem. Commonly used kernel functions include Gaussian kernel (with parameter γ)

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \tag{8}$$

and polynomial kernel (with parameter d)

$$K(x_i, x) = (\langle x_i, x \rangle + 1)^d. \tag{9}$$

When k (> 2) classes are concerned, two basic methods are available to convert a binary classification to k-class classification using SVM. In the "one against all" approach, k binary SVM are trained between each class and the remaining classes with the classification labeled to the class with maximum discriminant function. In the "one against one" approach, $k(k - 1)/2$ binary SVM are trained between each pair of classes with the classification labeled to the class with maximum votes.

The probabilistic estimates of SVM outputs are necessary if SVM is used to make partial decisions and the classification outputs need to be combined with other evidence (Platt, 2000). For binary SVM, a sigmoid function can be fitted to the discriminant function (Platt, 2000). For multi-class SVM, soft-max combination can be applied to "one against all" SVM outputs (Duan et al., 2003), and pairwise coupling of binary probabilistic estimates can be applied to "one against one" SVM outputs (Wu et al., 2004).

2.2. Markov Random Fields

Markov Random Fields (MRF) are commonly used probabilistic models for image analysis (Dubes & Jain, 1989; Li, 2001). The basic idea of MRF is to model the contextual correlation among image pixels in terms of conditional prior probabilities of individual pixels given their neighboring pixels. Based on Hammersley–Clifford theorem, the joint distribution of these conditional prior probabilities modeled by MRFs is equivalent to a joint prior probability characterized by a Gibbs distribution (Li, 2001). This MRFs–Gibbs equivalence allows us to model the complex global contextual relationship of an entire image by using MRFs of local pixel neighborhoods,

which make MRF computationally tractable and as a very popular contextual model.

Given a neighborhood system N (Fig. 1a), the spatial neighbors for a given pixel i are defined as $N_S(i)$. Let $c(i)$, $c(N_S(i))$ denote the class of pixel i and class vector of its spatial neighbors $N_S(i)$ respectively. Then an MRF models the spatial dependence of neighboring pixels as a conditional prior probability of a pixel as:

$$P(c(i)|c(N_S(i))) = \frac{1}{Z}\exp\left\{-U_S(c(i), c(N_S(i)))\right\} \quad (10)$$

where Z is the normalizing constant; U_S is the spatial energy function. The energy function U_S can be characterized as:

$$U_S(c(i), c(N_S(i))) = \beta_S \sum_{j \in N_S(i)} -I(c(i), c(j)) \quad (11)$$

where β_S is a non-negative parameter controlling the spatial dependence; I is equal to 1 if $c(i)=c(j)$ and 0 otherwise.

When multi-temporal image series are observed, MRF can be extended to include temporal dependences of pixels by adding a temporal energy function (Melgani & Serpico, 2003; Solberg et al., 1996). Suppose two image series are observed at time T_1 and $T_2 (T_1 < T_2)$. Assuming a second-order neighborhood system N (Fig. 1b), the temporal neighbors for a given pixel i are defined as $N_T(i)$. At time T_1, let $c_1(i)$, $c_1(N_S(i))$, $c_2(N_T(i))$ denote the class of a pixel i at time T_1, class vector of its spatial neighbors $N_S(i)$ at time T_1, and class vector of its temporal neighbors $N_T(i)$ at time T_2, respectively. Notations for time T_2 can be similarly defined. Then the spatial–temporal dependence of a pixel i at time T_1 given its spatial–temporal neighboring pixels is modeled as a conditional prior distribution by characterizing mutual influences among pixels as

$$\begin{aligned} &P(c_1(i)|c_1(N_S(i)), c_2(N_T(i))) \\ &= \frac{1}{Z}\exp\{-U(c_1(i), c_1(N_S(i)), c_2(N_T(i)))\} \end{aligned} \quad (12)$$

where Z is the normalizing constant; U is the total spatial–temporal energy function. If space–time interaction is not assumed, the energy function U can be modeled as the linear combination of spatial component and temporal component as in Eq. (13):

$$\begin{aligned} &U(c_1(i), c_1(N_S(i)), c_2(N_T(i))) \\ &= U_S(c_1(i), c_1(N_S(i))) + U_T(c_1(i), c_2(N_T(i))) \end{aligned} \quad (13)$$

where the spatial energy function U_S is defined as before; the temporal energy function U_T can be characterized by temporal transition probabilities as:

$$U_T(c_1(i), c_2(N_T(i))) = \beta_T \sum_{j \in N_T(i)} -P(c_1(j)|c_2(j)) \quad (14)$$

where $P(c_1(j)|c_2(j))$ is the transition probability of $c_2(j) \Rightarrow c_1(j)$; β_T is a non-negative parameter controlling the temporal dependence. The spatial–temporal dependence of a pixel j at time T_2 given its spatial–temporal neighbors can be similarly defined as analogy to Eqs. (12) (13) and (14).

MRF theory provides a convenient way for modeling the spatial–temporal contextual information in terms of conditional prior probabilities. However, MRF itself is not a sufficient classification rule but a prior model. Based on Bayesian decision rule, the optimal classification is achieved when the classification corresponds to maximum a posteriori (MAP). To achieve optimal classifications, MRF are often used in conjunction with algorithms estimating class conditional probabilities through the MAP decision rule. This conjunction gives rise to the MAP-MRF framework, under which class conditional probabilities are estimated from the observed spectral observations and combined with the prior contextual probabilities via Bayes formula (Li, 2001).

3. Study area and data

Our study area is located in China Camp State Park (122° 29 W, 38°00 N) in Marin County, California. The park is about 600 ha in size, with moderate to steep topography and elevations ranging from sea level at San Francisco Bay to 290 m. Vegetation in the park is varied but is dominated by mixed hardwood forest. We located our 12 ha study area within largely continuous mixed hardwood forest canopy. Overstory trees in the area range in age from 100 to 200 years. Coast live (*Quercus*
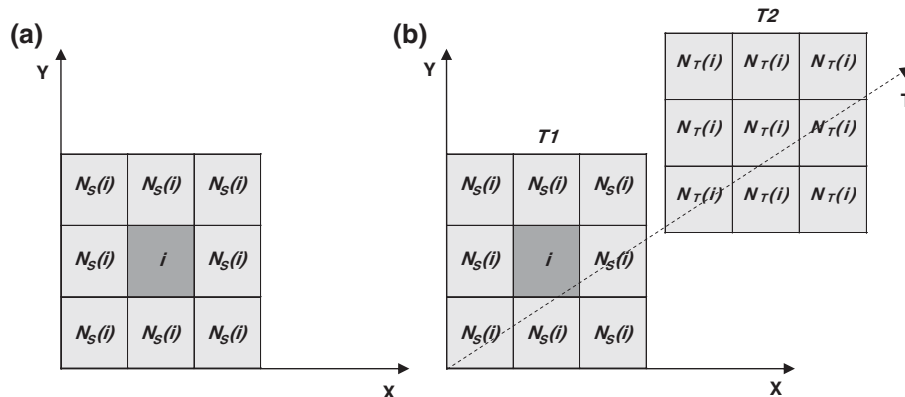

Fig. 1. MRF neighborhood systems: (a) spatial neighborhood; (b) spatial–temporal neighborhood. Axes X, Y are the spatial coordinates of pixels. Axis T is the temporal index of pixels.

agrifolia), black (*Quercus kelloggii)* and valley oaks (*Quercus lobata*) are common, and occur with mature madrone (*Arbutus menziesii*) and California bay (*Umbellularia californica*) trees. The understory is comprised of shrubs and small trees and vines, including manzanita, (*Arctostaphylos manzanita*), toyon (*Heteromeles arbutifolia*), hazel (*Corylus cornuta*), and buckeye (*Aesculus californica*). All of these plants with the exception of valley oak are hosts for *P. ramorum* (McPherson et al., 2005). We are targeting the two affected *Quercus* species (coast live oak and black oak) in this study, along with bare areas and the forest mosaic. These trees can be resolved using high-resolution imagery as their individual crowns range in diameter from 3 m to over 20 m. In addition, as diseased trees die, the entire crown changes dramatically from healthy green to brown over a short time period (in most cases) (Rizzo & Garbelotto, 2003).

Two ADAR 5500 (Airborne Data Acquisition and Registration) images acquired for the China Camp study area in March 30, 2000 and May 5, 2001 were used in this research. Imagery was acquired in spring because newly dead oak trees are most visible at this time, and to avoid mid-summer vegetation changes that might be confused with oak mortality; buckeye is a drought-deciduous species and can appear dead in summer. The ADAR 5500 imaging system is comprised of a SN4, 20 mm lens with four mounted cameras (spectral bands: band 1 (blue): 450–550 nm, band 2 (green): 520–610 nm, band 3 (red): 610–700 nm, band 4 (near infrared): 780–920 nm), flown at an average aircraft altitude of 2205 m. The cameras have a large field of view (typically about 35° from nadir) and collect high-spatial resolution data from relatively low altitudes with large aerial coverage. The average ground spatial resolution of our imagery is 1 m. Image acquisition and georeferencing were done by a private company (Positive Systems Inc. of Montana). The images were mosaicked, color balanced, topographically corrected, and then georeferenced to UTM projection using 16-cm ortho-photograph with the accuracy around 1 pixel. Atmospheric correction was not performed because it is not necessary to image classification (Song et al., 2000). The mosaicked image was clipped to our study area.

Since our major interest is to map the SOD dynamics across two years, we adopted the same classification scheme for both years. Specifically, three general classes to be identified from the two images are encoded as B (bare areas), D (dead oaks), and F (forest mosaics). Training and test samples were collected based on field visits, GPS data, and visual interpretation. All the dead stems were recorded in the field for the cause of death and geographic coordinates and identified on the images. Since tree death from reasons other than SOD is few, we assume all dead trees as SOD for simplicity.

4. Methods

4.1. Image-to-image registration

Accurate spatial registration of multi-temporal imagery is a necessity for our spatial–temporally explicit monitoring of SOD dynamics and subsequent analysis. The high spatial-resolution airborne ADAR imagery used in this paper had complex local deformation across different years due to the wide view angle, the terrain variation, the low flight height, and the effects of yaw, pitch, and roll (Devereux et al., 1990). These complicating factors present spatial challenges in aligning multi-temporal airborne imagery. Spatial registration was performed using manual control point extraction and global models, with pixel level accuracies. In this work, we improve upon these results with an automated method by combining area-based control point extraction methods with local geometric transformation models (Liu et al., in press). Specifically, the registration algorithm was implemented in two steps. First, an area-based method using correlation coefficient as similarity measure is applied to extract sufficient numbers of well-located control points. Second, the extracted control points are fed into a local transformation model, piecewise linear function, to register multi-temporal airborne images. More details are provided in Liu et al. (in press).

4.2. Specification of spatial energy function

The spatial energy function as introduced in Eq. (11) encourages neighboring pixels to be classified with the same labels and thus imposes a spatial smoothness effect on final classification. In general, this is a desirable prior model to avoid the "salt-and-pepper" effect inherent in non-contextual classifications. However, this general spatial energy function has limitations when spectral confusion exists between spatially smooth regions. Under such a circumstance, spatially neighboring pixels could be misclassified consistently so that the prior smoothness model does not provide any additional information to correct the spectral confusion. In particular in our study area, considerable spectral confusion exists between dead trees and bare areas. The spectral confusion resulted in misclassified pixels clustering in space. This made the general spatial energy function less useful.

In our spatial–temporal approach, the spatial energy function consists of two types of spatial relationships: spatial association and spatial exclusion. By spatial association, the algorithm favors two allocations: 1) spatially neighboring pixels in homogenous region being classified into the same class, and 2) spatially neighboring pixels in the boundary region being classified into ecologically compatible classes. This impacts a smoothing effect on the final classified map by screening out the isolated pixels in homogenous regions, yet keeping the boundary of different ground cover types in boundary region. With spatial exclusion, the algorithm penalizes the allocation that neighboring pixels are classified into ecologically incompatible classes. Specifically, the spatial energy function involved in MRF is specified in Eq. (15)

$$U_S(c(i), c(N_S(i))) = \beta_A \sum_{j \in N_S(i)} A(c(i), c(j)) + \beta_E \sum_{j \in N_S(i)} E(c(i), c(j)) \quad (15)$$

where β_A, β_E are parameters associated with spatial association and spatial exclusion; $A(c(i), c(j))$ is the association factor

between class $c(i)$ and class $c(j)$. $A(c(i), c(j))$ and $E(c(i), c(j))$ are defined as in Eqs. (16) and (17) respectively:

$$A(c(i), c(j)) = \begin{cases} -1, & \text{if} \quad c(i) = c(j) \\ 0, & \text{if} \quad c(i) \neq c(j) \end{cases} \qquad (16)$$

$$\begin{aligned} &E(c(i), c(j)) \\ &= \begin{cases} 1, & \text{if} \quad c(i) \text{ is incompatible with } c(j) \\ 0, & \text{if} \quad c(i) \text{ is compatible with } c(j) \end{cases}. \qquad (17) \end{aligned}$$

In the case of SOD, the occurrence of dead trees is associated with suitable forest environment because the spread of the disease is related to the distance from foliar hosts to target trees (Kelly & Meentemeyer, 2002; Rizzo & Garbelotto, 2003). Therefore, if pixels classified as dead trees are isolated from the forest mosaic by bare areas, they are assumed to be ecologically incompatible. Since dead trees are often overpredicted from bare areas, if we penalize the classification of dead trees in the neighborhood of bare areas through the spatial exclusion between dead trees and bare areas, we can correct the commission errors of dead trees from bare areas. In doing so, we define $c(i)=D$ is incompatible with $c(j)=B$ for $j \in N_S(i)$. The specific configurations involved in Eq. (15) are summarized in Table 1 (a).

4.3. Specification of temporal energy function

The temporal energy function defined in Eq. (14) can be understood as imposing a prior spatial smoothness indirectly from temporal neighbors. As the identity function suggests, spatial neighbors contribute to the spatial energy function in Eq. (11) in a crisp or discrete sense. In contrast, the temporal neighbors contribute to the energy function in a probabilistic

Table 1
Configurations of spatial and temporal energy functions

(a) Spatial energy function for both years

$c(j)$	B	D	F
$c(i)$			
B	$-\beta_A$	0	0
D	β_E	$-\beta_A$	0
F	0	0	$-\beta_A$

(b) Temporal energy function for year 2000

2001	B	D	F
2000			
B	$-\gamma_C P(B \Rightarrow B)$	0	$-\gamma_C P(B \Rightarrow F)$
D	γ_E	$-\gamma_C P(D \Rightarrow D)$	γ_E
F	$-\gamma_C P(F \Rightarrow B)$	$-\gamma_C P(F \Rightarrow D)$	$-\gamma_C P(F \Rightarrow F)$

(c) Temporal energy function for year 2001

2000	B	D	F
2001			
B	$-\gamma_C P(B \Rightarrow B)$	0	$-\gamma_C P(F \Rightarrow B)$
D	γ_E	$-\gamma_C P(D \Rightarrow D)$	$-\gamma_C P(F \Rightarrow D)$

sense, which is specified by transition probabilities $P(c_1(j)=c_1| c_2(j)=c_2)$. The larger $P(c_1(j)=c_1|c_2(j)=c_2)$, the more probable that pixel j at T_1 will be labeled j given pixel j at T_2 being labeled as c_2; and thus the more supports to pixel i at T_1 being labeled as c_1 is transferred indirectly from $c_2(j)$ through $c_1(j)$. Similarly, as there are limitations of spatial energy function to spectral confusion, the general temporal energy function in Eq. (14) also has limitations when the following two conditions are met: 1) spectral confusion exists between two classes c_1 and c_2 at T_1; and 2) c_1 or c_2 has prohibitive transition probability to another class c_3 at T_2. The limitation is originated from the transition probability model. Since the probability of any event is always greater than or equal to zero, when the prohibitive transition $c_1 \Rightarrow c_3$ or $c_2 \Rightarrow c_3$ happens, the probability of this transition is set to zero in Eq. (14). As a result, the prohibitive transition will have zero contributions to the final energy function, and only the prior spatial smoothness is posed on the final classification. The spectral confusion still remains and the prohibitive transition could appear in the final classifications. However, if the prohibitive transition were penalized in terms of negative contribution, it is possible to be used as complimentary information to solve the spectral confusion.

In our proposed spatial–temporal approach, the temporal energy function is characterized by two components: temporal relation and temporal exclusion. Temporal relation is modeled by a transition probability matrix which defines the probability of one pixel belonging to one land cover type at time T_1 given that it belongs to another cover type at time T_2. By temporal relation, the algorithm favors the allocation: 1) temporal neighboring pixels being classified into the same class if no change happens, and 2) temporal neighboring pixels being classified into ecologically compatible classes if a change happens. Since land cover change also satisfies spatial dependence, temporal relation will have a smoothing effect on the final change map by screening out the isolated change pixels in homogenous regions, yet keeping the boundary of different ground cover types in heterogeneous regions. By temporal exclusion, the algorithm penalizes the allocation: temporal neighboring pixels being classified into ecologically incompatible classes. This will discourage prohibitive transitions in the final classification map. Specifically, the temporal energy function in MRF model is formulated in Eq. (18):

$$\begin{aligned} U_T(c_1(i), c_2(N_T(i))) = &\gamma_C \sum_{j \in N_T(i)} -P(c_1(i)|c_2(j)) \\ &+ \gamma_E \sum_{j \in N_T(i)} I(c_1(i) \not\Rightarrow c_2(j)) \end{aligned} \qquad (18)$$

where γ_C, γ_E are parameters associated with temporal relation and temporal exclusion; $P(c_1(i)|c_2(j))$ is the temporal transition probability from class $c_1(i)$ to class $c_2(ij)$; $I(c_1(i)) \not\Rightarrow c_2(j))$ is 1 if the transition from $c_1(i)$ to $c_2(j)$ is prohibitive and 0 otherwise.

Based on our field investigation, no dead trees were removed in the study area during our time frame so dead trees at T_1 can be assumed to be dead at T_2. If pixels classified as dead trees at T_1 are classified as forest or bare areas at T_2, they are assumed to be

ecologically incompatible and the corresponding classification should be prohibited. Moreover, it is clear that bare areas at T_1 could not become dead trees at T_2. Therefore, there are three types of prohibitive transitions: (1) $B \Rightarrow D$, (2) $D \Rightarrow B$, and (3) $D \Rightarrow F$. By enforcing temporal exclusion for the three prohibitive transitions, we can correct the commission errors of dead trees from deciduous oaks at $T1$ and the spectral confusion between bare areas and dead oaks at T_1 and T_2. The specific configurations involved in Eq. (18) are summarized in Table 1 (b) and (c).

4.4. Spatial–temporal classification

The spatial and temporal energy functions specified in Sections 4.2 and 4.3 are the key parts of the MRF models. In the following, we develop our spatial–temporal classification algorithm based on these elements. The basic idea of the proposed algorithm is to simultaneously classify multi-temporal imagery by integrating the contributions from spectral observation of 4-spectral bands in ADAR images and spatial–temporal information among image pixels. Firstly, the spectral observation is the direct information source and the major evidence for classification in this algorithm. The most common statistical model used in spectral observation is that the class conditional probability (also called likelihood) of the observed spectral data is assumed to have a Gaussian distribution. Under such an assumption, maximum likelihood classification (MLC) algorithm is derived if no proper prior models are adopted. However, normal distribution assumption is sometimes inadequate to account for the possible multi-mode or skewed distribution. Alternatively, we propose to estimate the spectral likelihood using SVM for its free assumption and ability to find complex classification boundaries with good generalization performance as demonstrated by Liu et al. (2005). Secondly, the

spatial–temporal information serves two purposes in the proposed algorithm: 1) it provides complimentary information to spectral observation so that spectral confusion can be solved; and 2) it enforces constraints and regulations to adjust the spectral evidence so that the final classification could satisfy spatial–temporal dependence. In doing so, we model the spatial–temporal information in terms of conditional priors using MRF, within which the spatial and temporal dependence can be easily transformed into proper energy functions. Thirdly, the likelihood estimates of spectral observation from SVM and conditional spatial–temporal priors from MRF are integrated into posterior estimates by Bayes rule. In this manner, the spatial–temporally explicit classification algorithm combines the strength of SVM in spectral observation and effectiveness of MRF in spatial–temporal information, and the spectral, spatial, and temporal components of multi-temporal imagery are effectively unified.

Mathematically, the MAP-MRF framework is adopted to integrate spectral contributions modeled by SVM and spatial–temporal contributions modeled by MRF along the following lines. Assuming that a second order neighborhood system (Fig. 1b) and conditional independence of the spectral probability distribution, the posterior probability of pixel i belonging to class $c_1(i)$ given the observed spectral value $d_1(i)$ and its spatial–temporal neighbors ($c_1(N_S(i))$ and $c_2(N_T(i))$) is formulated as:

$$
\begin{aligned}
P(c_1(i)|d_1(i), c_1(N_S(i)), c_2(N_T(i))) &\propto P(d_1(i)|c_1(i)) \\
&\times P(c_1(i)|c_1(N_S(i)), c_2(N_T(i))) \\
&= \frac{1}{Z}\exp\{-[U_{\text{Spectral}} + U_{\text{Contextual}}]\}
\end{aligned} \tag{19}
$$

where U_{Spectral} is the energy function from the observed spectral data and $U_{\text{Spectral}} = -\ln(P(d_1(i)|c_1(i)))$, in which $P(d_1(i)|c_1(i))$ is
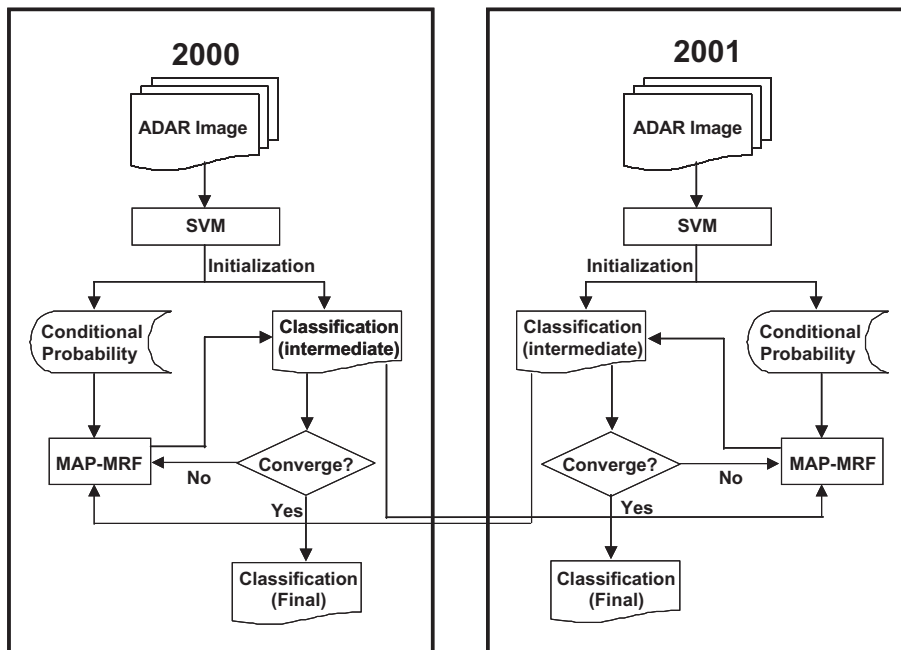

Fig. 2. The flowchart of the spatial–temporal explicit algorithm.

estimated by pairwise coupling of probabilistic estimates from "one against one" SVM outputs (Wu et al., 2004); and $U_{\text{Contextual}} = U_S + U_T$ is the energy function from spatial–temporal contextual information with U_S defined in Eq. (15) and U_T defined in Eq. (18). Since Z is constant for all classes, the MAP solution in Eq. (19) leads to the minimization of the sum of two energy functions as in Eq. (20).

$$c_i^*(i) = \arg \underset{c}{\text{Min}} \{U_{\text{Spectral}} + U_{\text{Contextual}}\}. \qquad (20)$$

Theoretically, the optimal solution of Eq. (20) should be obtained by simultaneously minimizing the sum of two energy functions for all pixels. Global optimization algorithms such as Simulated Annealing (SA) (Geman & Geman, 1984) and Maximum a Posterior Margin (MPM) (Marroquin et al., 1987) are available to get the optimal solution. However, these algorithms are computationally demanding in general. Instead, a simple and computationally less demanding algorithm called Iterative Conditional Modes (ICM) is used to obtain a sub-optimal solution (Besag, 1986). The spatial smoothing weight and temporal transition weight are determined by cross-validation. The transition probabilities are iteratively estimated from the classified images. The spatial and temporal exclusion penalty terms are heuristically determined as long as they are large enough to penalize the spatially and temporally incompatible classification.

The implementation of the proposed spatial–temporally explicit classification algorithm is illustrated in Fig. 2. The algorithm consists of the following two stages: initialization and iteration. In the initialization stage, SVM is trained with the spectral observations to allocate initial class labels and estimate pixel-wise class conditional probabilities for each individual image. In the iteration stage, first, MRF are used to model spatial–temporal contextual prior probabilities of images based on the intermediate classification from the previous iteration; next, the conditional probability estimates from SVM and prior probabilities from MRF are converted into appropriate energy functions and unified under the MAP–MRF framework; following that, the classification is updated based on the combination of spectral class conditional probability and spatial–temporal contextual prior probabilities; and finally, the iteration proceeds until the predefined convergence rate is met.

4.5. Accuracy assessment

Accuracy assessments were performed for the classifications based on the ground truth data reserved for this purpose. The ground truth data for accuracy assessments were randomly generated based on visual interpretation, field visits, and GPS data. These testing data are representative of the full range of the classes and are independent of the training data. The amount of the testing data for each class can be found in Tables 3 and 4. Error matrices or confusion matrices were generated for all the classifications with different algorithms. Producer's accuracy (probability of a reference pixel being correctly classified), user's accuracy (probability that a pixel classified on the final

product represents that category on the ground), overall accuracy, and a Kappa coefficient were derived from each error matrix. The significances in the difference between different classifications were also evaluated based on the Kappa coefficients and their corresponding variance estimates.

5. Results

5.1. Image-to-image registration

The image of 2001 was registered to that of 2000 using a piecewise linear transformation model with the control points automatically extracted from an area-based method. The control points were evenly arranged on regular grids with a grid size of 30 pixels. In total, 100 control points were generated automatically. The registration error was estimated based on 15 independent control points as check points selected manually. The RMSE (root mean squared errors) is 0.43-pixel, which is significantly better than the previous results from second polynomial transformation model (larger than 2-pixel).

5.2. Non-contextual classification

The initial results of non-contextual classifications are reported here for two purposes: 1) they are the initializations of subsequent spatial–temporal classifications; and 2) they will be used as benchmark for comparison with our spatial–temporal methods. For each year, we trained a SVM classifier with a Gaussian kernel function using data from all four bands. This classifier involved two parameters: the penalty term C and the Gaussian smoothing parameter γ. We found the optimal values (Table 2) for the two parameters by using a 5-fold cross-validation method. The initial non-contextual classification maps were shown in Fig. 4 (a)–(b). For comparison, we also implemented MLC for non-contextual classification. The classification maps were shown in Fig. 3 (a)–(b). The accuracy assessments of all the classifications are presented in terms of several measures based on the independent test samples (Table 3). The comparison of the classification maps between MLC and SVM is reported in Table 5.

For the image of 2000, the non-contextual MLC resulted in moderate accuracy (overall: 80.1%, Table 3) with considerable "speckles". However, as mentioned in the introduction, the user's accuracy of dead oaks (47.7%, Table 3) is low due to two types of commissions: 1) spectral confusion with bare areas, and

Table 2
Parameters for SVM classifiers and MRF models

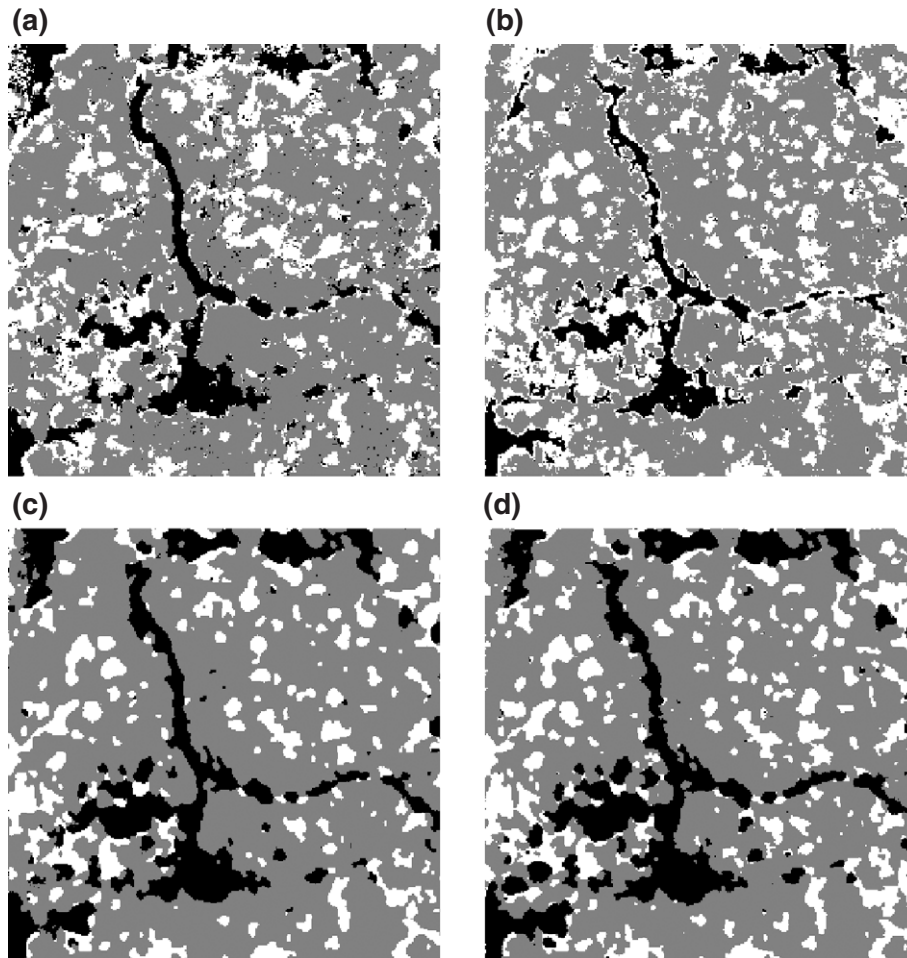|	Parameter	Image of 2000	Image of 2001
SVM	C	100	1000
	γ	50	30
MRF	β_A	0.85	0.7
	β_E	10	10
	γ_A	0.6	0.5
	γ_E	10	10

Fig. 3. Classified images using MLC based methods where B (Bare areas), D (Dead oaks), and F (Forest mosaics) are represented in black, white and gray respectively: (a) non-contextual MLC (2000); (b) non-contextual MLC (2001); (c) spatial–temporal MLC (2000); and (d) spatial–temporal MLC (2001).

2) confusion with deciduous oak caused by seasonality. As shown in Fig. 3 (a)–(b), the commissions from bare areas are mostly distributed along the edges between forest and bare areas, where grass and other understory vegetation and bare areas are mixed to appear spectrally more similar to dead crowns than the other two classes. The commissions from deciduous oaks are mainly located at the central north part of the study area, which can be observed by comparing with the classification map of 2001. Non-contextual SVM performed moderately well (overall: 87.4%; user's accuracy for dead: 60.8%, Table 3) producing fewer speckles. This result was significantly better than MLC (Table 5). It had less commission errors from bare areas, but there was still considerable confusion between dead and deciduous forest (Fig. 4 (a)–(b)). The better performance of SVM may be attributed to its good generalization ability and non-parametric nature of the statistical learning rule whereas the normal distribution assumption of MLC may not be satisfied for forest and bare areas, which exhibit multi-mode distribution. For image of 2001, similar results were obtained except that commission errors of dead oaks were mainly from bare areas because the deciduous oaks were not fully leafed out at the time of imaging and were successfully classified to forest.

The classified transitions from year 2000 to year 2001 were identified by overlaying the two classified images from non-contextual algorithms (Fig. 5 (a), (c)). With three land cover types {B, D, F}, there are six types of possible land cover transitions from one type in 2000 to another in 2001: {BD, BF, DB, DF, FB, FD}. Among the six transitions, three transitions {BD, DB, DF} are prohibitive. On the classified transition map from non-contextual MLC (Fig. 5-(a)), the three prohibitive transitions were visible across the whole image due to the misclassification. In particular, considerable DF transition (in yellow color) was caused by the spectral confusion between deciduous oaks and dead crowns. Comparatively, fewer prohibitive transitions were shown on the classified transition map from non-contextual SVM (Fig. 5-(c)).

5.3. Spatial–temporal classification

The spatial–temporal classification algorithms started on the basis of the initial non-contextual classifications as reported above. The spectral conditional probabilities from SVM or MLC were integrated with spatial–temporal contextual information under the framework of MAP–MRF, which is characterized by parameters listed in Table 2. The MAP–

Table 3
Accuracy matrices for non-contextual classifications

			Reference				
		Classified	Bare	Dead	Forest		User's
Non-Contextual MLC	Image of 2000	Bare	893	5	91	989	90.3%
		Dead	186	865	761	1812	47.7%
		Forest	0	3	2441	2444	99.8%
			1079	873	3293	5245	
		Producer's	82.8%	99.1%	74.1%	Overall	80.1%
						Khat	0.67
						Z	**0.000079**
	Image of 2001	Bare	786	12	2	800	98.3%
		Dead	430	1152	746	2328	49.5%
		Forest	2	0	2450	2452	99.9%
			1218	1164	3198	5580	
		Producer's	64.5%	98.9%	76.6%	Overall	78.6%
						Khat	0.66
						Z	**0.000071**
Non-Contextual SVM	Image of 2000	Bare	920	11	73	1004	91.6%
		Dead	142	845	403	1390	60.8%
		Forest	17	17	2817	2851	98.8%
			1079	873	3293	5245	
		Producer's	85.3%	96.8%	85.6%	Overall	87.4%
						Khat	0.78
						Z	**0.000061**
	Image of 2001	Bare	884	11	6	901	98.1%
		Dead	322	1150	402	1874	61.4%
		Forest	12	3	2790	2805	99.5%
			1218	1164	3198	5580	
		Producer's	72.6%	98.8%	87.2%	Overall	86.5%
						Khat	0.78
						Z	**0.000053**

Bold numbers indicate results are significantly better than random chance at the 95% confidence level.

MRF was solved iteratively using the ICM algorithm. As Fig. 6 illustrated, ICM converged with a convergence rate below threshold 0.005% in 7 iterations. The final spatial–temporal classification maps were shown in Figs. 3 (c)–(d) and 4 (c)–(d). The accuracy assessments of all the classifications are summarized in Table 4. The comparison of the classification maps between contextual MLC and contextual SVM is reported in Table 5.

For the image of 2000, the spatial–temporal contextual classification based on both MLC and SVM improved upon the initial non-contextual classification (MLC: overall accuracy improved from 80.1% to 87.2%; user's accuracy of dead improved from 47.7% to 58.0%. SVM: overall accuracy improved from 87.4% to 95.9%; user's accuracy of dead improved from 60.8% to 89.8%, Table 4). The improvements were significant (Table 5). The improvements were also illustrated by the reduction of speckles due to spatial smoothing (Figs. 3 and 4). With the incorporation of spatial exclusion, the commissions from bare areas along the edges between forest and bare areas were dramatically reduced. The commission errors from deciduous oaks were also corrected with the temporal exclusion from 2001 image. In comparison with the MLC based spatial–temporal classification, SVM based classifications performed significantly better (Table 4) with less commission errors from bare areas and deciduous forest. For the 2001 image, similar results were observed.

By overlaying the two classified images, we got the classified transition maps from year 2000 to year 2001 with the incorporation of spatial–temporal information (Fig. 5 (b), (d)). On the classified transition maps from both spatial–temporal MLC (Fig. 5-(b)) and spatial–temporal SVM (Fig. 5-(d)), the three prohibitive transitions have been dramatically reduced compared to their non-contextual counterparts. Specifically, the bare-to-dead (BD: in red) transition was nearly invisible. The dead-to-bare (DB: in blue) and dead-to-forest (DF: in yellow) transitions were largely removed and converted to no change. These improvements are apparently due to the added temporal exclusions.

6. Discussion

In this paper, we developed a new spatial–temporally explicit classification algorithm using ADAR imagery from two dates in an attempt to improve upon non-contextual methods for forest disease monitoring. The classification errors caused by spectral confusions between dead oaks with bare areas and healthy deciduous oaks (in year 2000) are largely reduced with the incorporation of the spatial–temporal information in the classification. This is clearly evident in the correction in ecologically incompatible allocation of dead trees at the bare areas shown in Fig. 4 (c)–(d) and in the reduction in less likely transitions bare-to-dead (BD: in red),
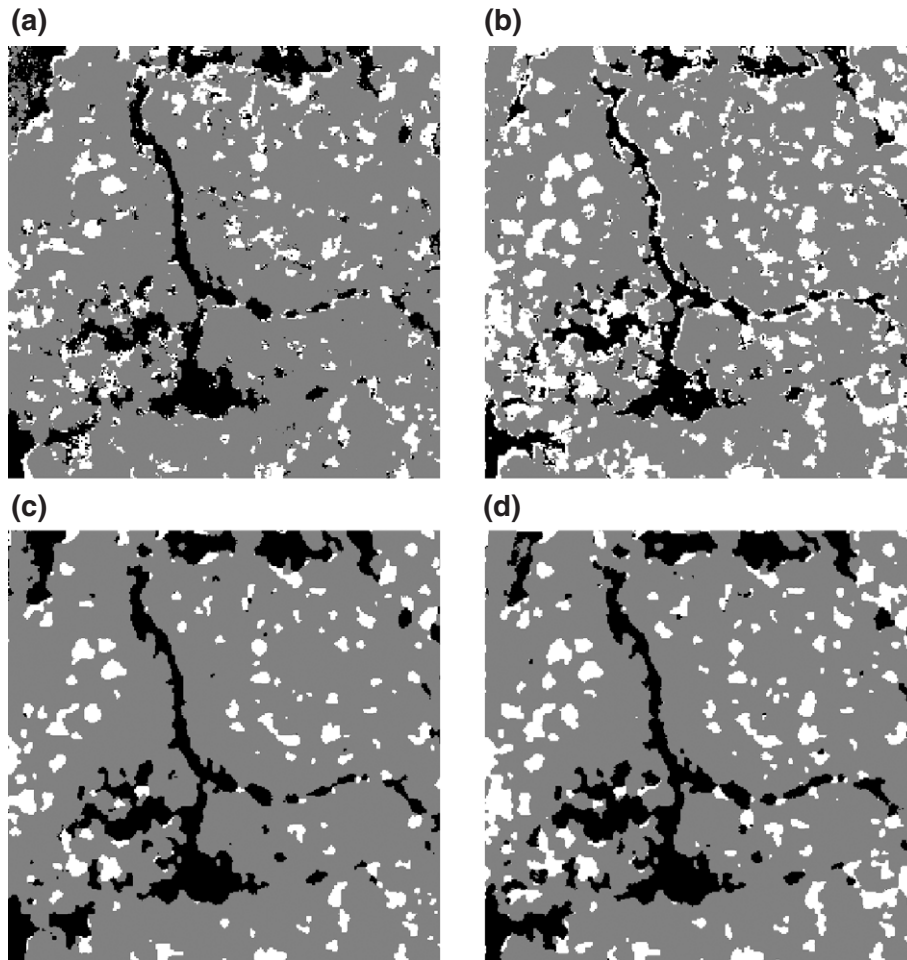
Fig. 4. Classified images using SVM based methods where B (Bare areas), D (Dead oaks), and F (Forest mosaics) are represented in black, white and gray respectively: (a) non-contextual SVM (2000); (b) non-contextual SVM (2001); (c) spatial–temporal SVM (2000); and (d) spatial–temporal SVM (2001).

dead-to-bare (DB: in blue), and dead-to-forest (DF: in yellow) shown in Fig. 5. The larger clumps of forest-to-dead (FD: in cyan) are our real target: these are clumped pixels that changed from healthy trees to dead crowns. These clumps can be used to further model the spatial pattern of disease through a forest stand. The remaining sparsely distributed dead-to-forest (DF: in yellow) and forest-to-dead (FD: in cyan) transitions are found in linear curves that correspond to canopy edges. There are three possible sources of these linear shapes: 1) errors in registration, 2) classification errors at the canopy edges where spectral uncertainty exists due to mixed pixel effects, and 3) morphological change of canopy shape. To get clearer transition map, post-processing such as spatial filtering could be applied to remove these spurious artifacts. Further improvement of the algorithm is needed to reduce these artifacts.

Spatial smoothing, as a general rule for spatial dependence, is widely used in traditional classification models for the use of spatial contextual information. However, while spatial smoothing can reduce "speckle" found in traditional pixel-based classification, it is unable to remove misclassifications created by ecologically incompatible but spatially smooth classification results. We thus introduced a spatial exclusion rule to correct those spatially smooth but ecologically incompatible classification. Spatial exclusion has been less frequently examined than has spatial smoothing in remote sensing contextual classification. Spatial exclusion is more domain-specific and can be used as a way to encode an expert's knowledge about the study area. From the results of this study, we demonstrated an example of the principles underlying the incorporation of spatial exclusion in the classification algorithms. More examples can be developed in different applications. For example, in land use classifica-tion, the spatial exclusion can be established based on the social and economic factors underlying different land use patterns. In vegetation species classification, the alliance relationship, ecological and topological knowledge can be used in characterizing spatial exclusion rules.

We also used temporal dependence to relate images from two dates. Our proposed algorithm can be generalized to larger image series. When three or more images are considered, one image can be linked with its past neighboring image and its future neighboring image if a Markov property is assumed. The classification maps are simultaneously updated across three or more images. The phenological correction used in our application just provides one example
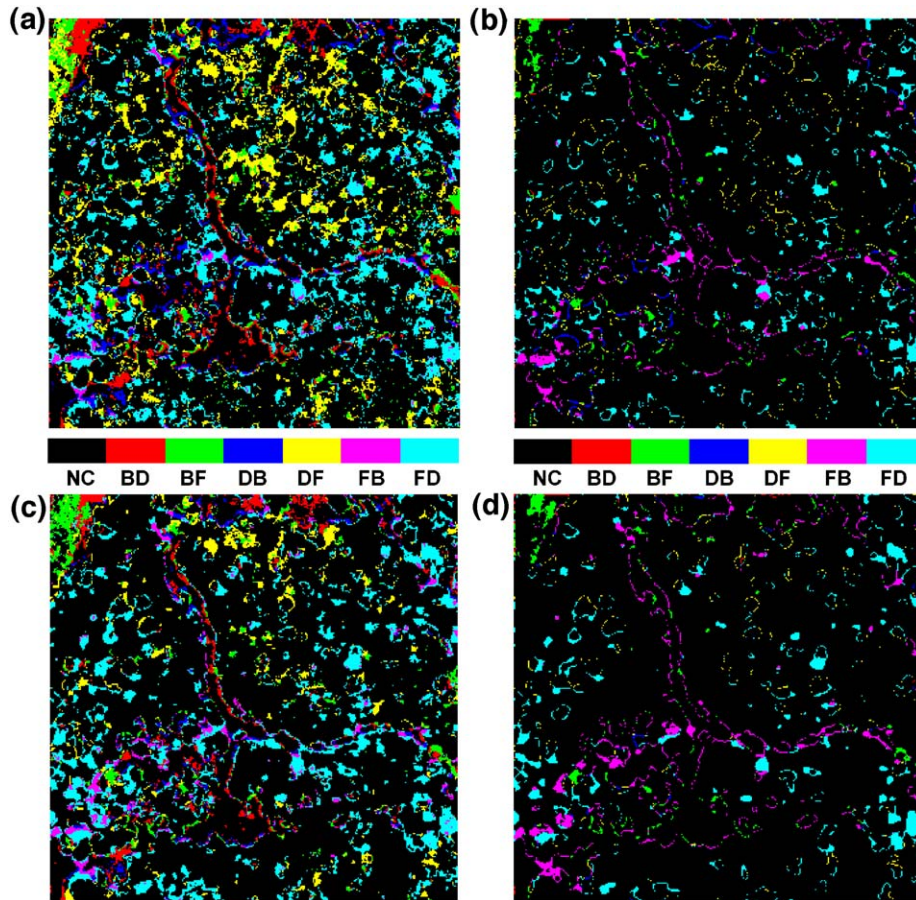
Fig. 5. The classified transitions between two years: (a) non-contextual MLC; (b) spatial–temporal MLC; (c) non-contextual SVM; and (d) spatial–temporal SVM. Notation: NC (No Change), BD (Bare to Dead), BF (Bare to Forest), DB (Dead to Bare), DF (Dead to Forest), FB (Forest to Bare), FD (Forest to Dead).

demonstrating how temporal information can be used to solve cases of spectral confusion. More specific application can be designed to encode temporal relations into specific allocation rules. Specifically, changes can be related to land use history,

Fig. 6. The convergence curves of ICM initialized by MLC and SVM for both years.

fire inventory and vegetation phenology and we can use these ancillary information to establish temporal relation among different image series. For example, in Liu et al. (2005), a fire perimeter layer is used to link the land cover transition before and after fires; in Aurdal et al. (2005), a phenology model based on Markov chain is used to integrate multi-temporal TM imagery.

Overall, our ability to extract in an automated fashion the tree canopy that changes from healthy to dead, while excluding other less likely transitions will contribute to further spatial–temporal analysis of disease dynamics and landscape disease modeling. This work using a spatial–temporally explicit algorithm in a forest health-monitoring context has broader applicability across different applications using multi-temporal remote sensing imagery. The overall methodology presented in this paper provides a general framework on the use of spatial–temporal information from multi-temporal remote sensing data. As such, the new method is not limited to the specific data and classification scheme as used in this work. It is possible to use the proposed spatial–temporal framework to design various multi-temporal monitoring projects with different classification schemes (e.g. more classes) and different types (e.g. hyperspectral imagery) of imagery. For example, a similar approach has been demonstrated in Liu et al. (2005) for land cover classification with
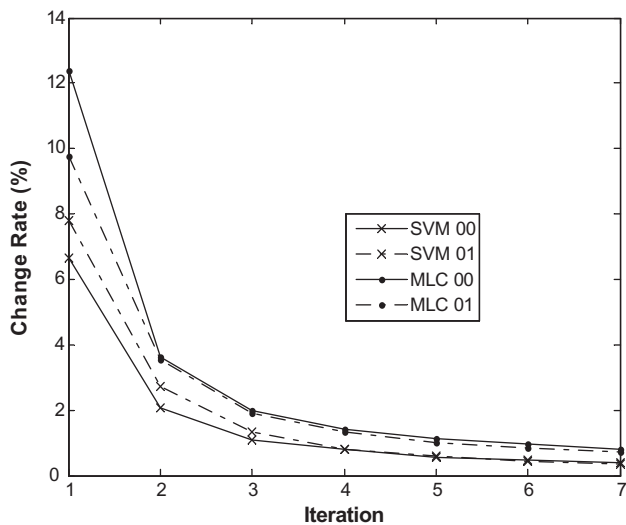
Table 4
Accuracy matrices for spatial–temporal classifications

		Reference				
		Classified	Bare	Dead	Forest	User's
Spatial–temporal MLC	Image of 2000	Bare	1003	12	28	1043	96.2%
		Dead	68	852	548	1468	58.0%
		Forest	8	9	2717	2734	99.4%
			1079	873	3293	5245	
		Producer's	93.0%	97.6%	82.5%	Overall	87.2%
						Khat	0.78
						Z	**0.000061**
	Image of 2001	Bare	1126	31	40	1197	94.1%
		Dead	86	1126	627	1839	61.2%
		Forest	6	7	2531	2544	99.5%
			1218	1164	3198	5580	
		Producer's	92.5%	96.7%	79.1%	Overall	85.7%
						Khat	0.77
						Z	**0.000056**
Spatial–temporal SVM	Image of 2000	Bare	1006	18	27	1051	95.7%
		Dead	28	821	65	914	89.8%
		Forest	45	34	3201	3280	97.6%
			1079	873	3293	5245	
		Producer's	93.2%	94.0%	97.2%	Overall	95.9%
						Khat	0.92
						Z	**0.000026**
	Image of 2001	Bare	1108	10	28	1146	96.7%
		Dead	61	1132	136	1329	85.2%
		Forest	49	22	3034	3105	97.7%
			1218	1164	3198	5580	
		Producer's	91.0%	97.3%	94.9%	Overall	94.5%
						Khat	0.91
						Z	**0.000027**

Bold numbers indicate results are significantly better than random chance at the 95% confidence level.

nine classes using 30 m TM imagery. However, the spatial–temporal information might vary its value in different applications with different spatial resolution and temporal frequency of images. For example, high spatial resolution multi-date images might benefit more from the spatial information than coarse resolution image because the spatial dependence of pixels at high spatial resolution tends to be stronger than at low spatial resolution. On the other hand, it is easier to obtain higher temporal frequency data with coarser spatial resolution and thus the spatial–temporal classifier might benefit more from the temporal sources of information with coarser spatial resolution of data.

Table 5
Pairwise comparison of error matrices for all classifiers

Pairwise comparison					
		MLC	SVM	SP–MLC	SP–SVM
Image of 2000	MLC	0.00			
	SVM	**9.01**	0.00		
	SP–MLC	**9.05**	0.05	0.00	
	SP–SVM	**24.33**	**15.30**	**15.25**	0.00
Image of 2001	MLC	0.00			
	SVM	**10.40**	0.00		
	SP–MLC	**9.76**	0.56	0.00	
	SP–SVM	**24.82**	**14.52**	**14.90**	0.00

Bold results are significant at the 95% confidence level.

7. Conclusions

Our results indicate that a spatial–temporal classification algorithm that explicitly integrates spectral, spatial and temporal information in multi-temporal high-spatial resolution images can achieve significant improvements over non-contextual classifications, and specific to our case, reduce the commission errors produced from other land types being classified as dead trees. In addition, our results also show that SVM performed better in the processing of spectral data than MLC, and provided better classification initialization for our spatial–temporal classification. We have three specific conclusions. First, SVM are promising algorithms for spectral classification and can be improved with spatial–temporal information. Second, MRF are efficient probabilistic models for the analysis of spatial and temporal contextual information. And third, the combination of SVM and MRF in a MAP–MRF framework unifies the strengths of two algorithms and leads to an improved integration of the spectral, spatial, and temporal components of multi-temporal remote sensing images. The improved accuracies generated by the addition of spatial–temporal contextual information prove the importance of spatial–temporal modeling to multi-temporal remote sensing.

Acknowledgements

References

Asner, G. P., Keller, M., Pereira, J. R., Zweede, J. C., & Silva, J. N. M. (2004). Canopy damage and recovery after selective logging in Amazonia: Field and satellite. *Ecological Applications*, *14*, S280–S289.

Aurdal, L., Husedy, R. B., Eikvil, L., Solberg, R., Vikhamar, D., & Solberg, A. (2005). Use of hidden Markov models and phenology for multitemporal satellite image classification: Applications to mountain vegetation classification. *Proceedings of the IEEE third international workshop on the analysis of multitemporal remote sensing images* (pp. 220–224).

Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, B*, *48*, 259–302.

Boyer, M., Miller, J., Belanger, M., & Hare, E. (1988). Senescence and spectral reflectance in leaves of northern pin oak (*Quercus palustris Muenchh.*). *Remote Sensing of Environment*, *25*, 71–87.

Bruzzone, L., & Smits, P. C. (2001). *Analysis of multi-temporal remote sensing images.* New Jersey: World Scientific.

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining*, *2*(2).

Clark, D. B., Castro, C. S., Alvarado, L. D. A., & Read, J. M. (2004a) Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data. *Ecological Letters*, *7*, 52–59.

Clark, D. B., Read, J. M., Clark, M. L., Cruz, A. M., Dotti, M. F., & Clark, D. A. (2004b) Application of 1-m and 4-m resolution satellite data to ecological studies of tropical rain forests. *Ecological Applications*, *14*, 61–74.

Cristianini, N., & Shawe-Taylor, J. (2001). *An introduction to support vector machines.* Cambridge, UK: Cambridge University Press.

Devereux, B. J., Fuller, R. M., Carter, L., & Parsell, R. J. (1990). Geometric correction of airborne scanner imagery by matching Delaunay Triangles. *International Journal of Remote Sensing*, *11*(12), 2237–2251.

Duan, K., Keerthi, S. S., Chu, W., Shevade, S. K., & Poo, A. N. (2003). Multi-category classification by soft-max combination of binary classifiers. *The fourth international workshop on multiple classifier systems*.

Dubes, R. D., & Jain, A. K. (1989). Random field models in image analysis. *Journal of Applied Statistics*, *16*, 131–164.

Everitt, J., Escobar, D., Appel, D., Riggs, W., & Davis, M. (1999). Using airborne digital imagery for detecting oak wilt disease. *Plant Disease*, *83*(6), 502–505.

Garbelotto, M., Svihra, P., & Rizzo, D. (2001, Jan/Feb). Sudden oak death syndrome fells 3 oak species. *California Agriculture*, 9–19.

Geman, S., & Geman, D. (1984). Stochastic relaxation Gibbs distributions and the Bayesian restoration of the image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*(6), 721–741.

Gong, P., Mei, X., Biging, G. S., & Zhang, Z. (1999). Monitoring oak woodland change using digital photogrammetry. *Journal of Remote Sensing*, *3*(4), 285–289.

Gong, P., & Xu, B. (2003). Remote sensing of forests over time: Change types, methods, and opportunities. In M. Woulder, & S. E. Franklin (Eds.), *Remote sensing of forest environments: Concepts and case studies* (pp. 301–333). Amsterdam, Netherlands: Kluwer Press.

Holdenrieder, O., Pautasso, M., Weisberg, P., & Lonsdale, D. (2004). Tree diseases and landscape processes: The challenge of landscape pathology. *Trends In Ecology and Evolution*, *19*, 446–452.

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(40), 725–749.

Jeon, B., & Landgrebe, D. A. (1999). Decision fusion approach for multitemporal classification. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 12233–12270.

Kelly, N. M. (2002). Monitoring Sudden Oak Death in California using high-resolution imagery. *USDA-Forest Service General Technical Report PSW-GTR-184* (pp. 799–810).

Kelly, M., & Meentemeyer, R. K. (2002). Landscape dynamics of the spread of Sudden Oak Death. *Photogrammetric Engineering and Remote Sensing*, *68* (10), 1001–1009.

Kelly, M., Shaari, D., Guo, Q., & Liu, D. (2004a). A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by "sudden oak death". *Photogrammetric Engineering and Remote Sensing*, *70*, 1229–1239.

Kelly, M., Tuxen, K., & Kearns, F. (2004b).Geospatial informatics for management of a new forest disease: Sudden oak death. *Photogrammetric Engineering and Remote Sensing*, *70*, 1001–1004.

Key, T., Warner, T., McGraw, J., & Fajvan, M. (2001). A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sensing of Environment*, *75*, 100–112.

Li, S. Z. (2001). *Markov random field modeling in image analysis.* Springer.

Liu, D., Gong, P., Kelly, M., and Guo, Q., (in press). Automatic registration of airborne image with complex local deformation. *Photogrammetric Engineering and Remote Sensing*.

Liu, D., Kelly, M., & Gong, P. (2005). Classifying multi-temporal TM imagery using Markov random fields and Support Vector Machines. *Proceedings of the IEEE third international workshop on the analysis of multitemporal remote sensing images* (pp. 225–228).

Marroquin, J., Mitter, S., & Poggio, T. (1987). Probabilistic solution of illposed problems in computational vision. *Journal of the American Statistical Association*, *82*(397), 76–89.

McPherson, B. A., Mori, S. R., Wood, D. L., Storer, A. J, Svihra, P., Kelly, N. M., et al. (2005). Sudden oak death in California: Disease progression in oaks and tanoaks. *Forest Ecology and Management*, *213*, 71–89.

Melgani, F., & Serpico, S. B. (2003). A Markov random field approach to spatio-temporal contextual image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(11), 2478–2487.

Muchoney, D. M., & Haack, B. N. (1994). Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, *60* (10), 1243–1251.

Platt, J. (2000). *Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. Advances in large margin classifiers.* Cambridge, MA: MIT press.

Richards, J. A., & Jia, X. (1999). *Remote sensing digital image analysis: An introduction*, 3rd edition. New York: Springer-Verlag.

Ristaino, J. B., & Gumpertz, M. (2000). New frontiers in the study of dispersal and spatial analysis of epidemics caused by species in the genus *Phyophthora*. *Annual Review of Phytopathology*, *38*, 541–576.

Rizzo, D., & Garbelotto, M. (2003). Sudden oak death: Endangering California and Oregon forest ecosystems. *Frontiers in Ecology and the Environment*, *1* (5), 197–204.

Rizzo, D., Garbelotto, M., Davidson, J. M., Slaughter, G. W., & Koike, S. T. (2002). *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease*, *86* (3), 205–213.

Solberg, A. H. S., Taxt, T., & Jain, A. K. (1996). A Markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(1), 100–113.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2000). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, *75* (2), 224–230.

Vapnik, V. (1998). *Statistical learning theory.* Wiley.

Wu, T. F., Lin, C. J., & Weng, R. C. (2004). Probability estimates for multi-class classification by pairwise coulpling. *Journal of Machine Learning Research*, *5*, 975–1005.

Wulder, M. A., Hall, R. J., Coops, N. C., & Franklin, S. E. (2004). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, *54*, 511–521.