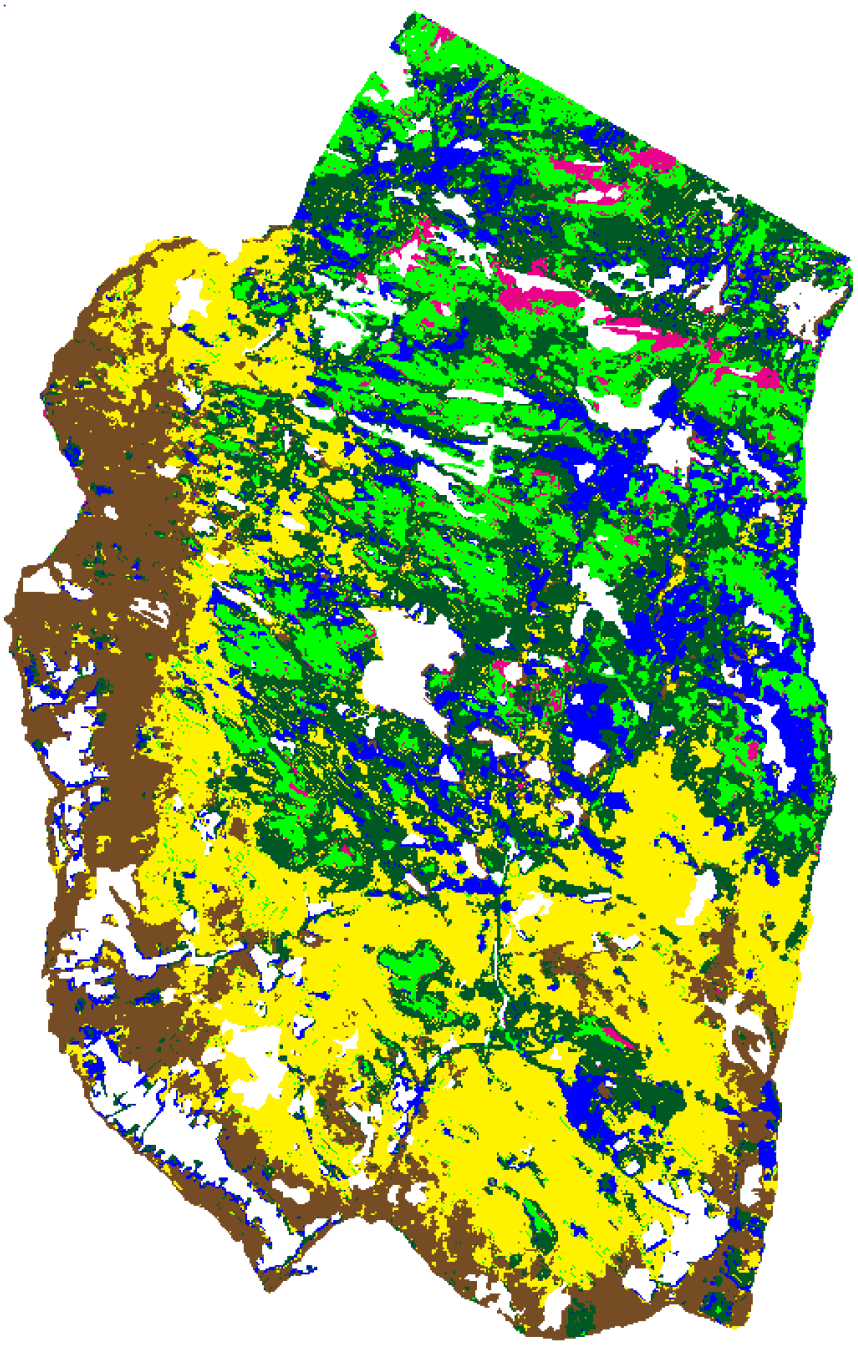
Classification of multitemporal satellite images using phenological models

Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NOTE



SAMBA/29/05

Lars Aurdal
Ragnar Bang Huseby
Dagrun Vikhamar
Line Eikvil
Anne Schistad Solberg
Rune Solberg

27th September 2005

Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NR Note

Title: Classification of multitemporal satellite images using phenological models

Date: 27th September 2005
Year: 2005
Note no.: SAMBA/29/05

Author: Lars Aurdal <lars.aurdal@nr.no>
Ragnar Bang Huseby
<ragnar.huseby@nr.no>
Dagrun Vikhamar
<dagrun.vikhamar@geo.uio.no>
Line Eikvil <line.eikvil@nr.no>
Anne Schistad Solberg <anne@ifi.uio.no>
Rune Solberg <rune.solberg@nr.no>

Abstract:

Keywords: multitemporal satellite image classification phenology model

Target group: Open

Availability: Open

Project: EO-Tools

Project no.: 830110

Research field: Remote sensing

No. of pages: 23

Norwegian Computing Center
Gaustadalléen 23, P.O. Box 114 Blindern, NO-0314 Oslo, Norway
Telephone: +47 2285 2500, telefax: +47 2269 7660, http://www.nr.no

Contents

1 Introduction

Optical satellite images are gaining an ever increasing importance as a tool for monitoring various aspects of the earths surface. In spite of the steadily rising quality of these images, the interpretation remains a challenging task. This is mainly so due to the complexity of the observed scenes. The earths surface, and its appearance when observed from space, is influenced by a large number of factors. Satellite images of the very same scene, even acquired close in time, can vary widely. This is so for many reasons. Possibly the most obvious cause for this variation are clouds and their shadows. Moving around in the atmosphere with the shifting winds, clouds can change the observability of a region to near zero in a very short time. Similar natural variations can occur on many scales, both spatial and temporal. Some of these are:

- Atmosphere temperature and humidity variations.
- Atmosphere pollutants, both natural (sand drifts, volcano smoke) and man made (industry).
- Seasonal variations in ground cover.
- The presence or not of snow.

These and many other factors make it very hard to interpret a satellite image. One possible remedy is that of increasing the resolution of the acquired images. Over the last years we have seen a spectacular increase in the quality of the acquired images, both in **spectral** and **spatial** resolution. Another remedy is that of increasing the **temporal** resolution of the images, that is, basing the interpretation of the satellite scene on a **sequence of satellite images densely sampled in time**. It is this approach that will be explored in the current project. In particular we aim at developing methods for classifying satellite images where classification is based on the **temporal evolution** of a class and not on its appearance in a single scene.

1.1 Aim of this project

The main aim of this project is the establishment of algorithms for classifying satellite images based on **temporal sequences** of satellite images of the same scene, thus the classification will be based on how different classes evolve in time, and not only on how they appear in a single scene.

1.2 Overview of method

The task of multi-temporal satellite image classification is, as we have already pointed out, a very complex task. The method we have chosen for solving this problem reflects this,- it is composed of a sequence of complicated steps all serving very specific purposes. Initially the problem is split in two separate problems, the problem of *image calibration and intercalibration* and the problem of *image classification*.

The image calibration problem consists of two steps, that of georeferencing the images and that of correcting them radiometrically. The georeferencing is necessary in order to establish the correspondence between satellite image pixels and physical positions on the earths surface. The radiometric correction consists in compensating for satellite sensor gain and offset, furthermore we subtract the average radiance measured over deep water in all channels. This will to a large degree compensate for atmospheric reflection. Geometric correction was performed on the image data used in this project, radiometric correction was not performed.

The image intercalibration is the subject of a separate report, see Huseby et al. (2005). This step consists in combining satellite images, possibly from several seasons, into one single time series. This is obviously not a straightforward procedure since the phenological evolution over a season can vary greatly from one year to another. It is e.g. quite possible that an early season in one year will produce late spring images that would match early summer images from another, late, season another year. This step is necessary in order to make data from different seasons comparable and in order to make it possible to combine such data into one single time series.

The image classification step is performed using a Hidden Markov Model (HMM) for the phenology of the ground cover. Basically, the ground cover is assumed to be in one of a series of discrete states. In each of these states the different ground cover classes will result in observables linked to the underlying state only in a statistical sense. The classification consists in choosing the class that best explains the observables.

2 Acknowledgments

This project was partially financed by the Research council of Norway. We gratefully acknowledge the Norwegian Institute of land inventory for having provided the digital vegetation maps used in the project. We also thank Professor Dag Klaveness of the Department of biology, the University of Oslo for interesting discussions on plant phenology.

2.1 Organization of this document

In the next section we give a brief overview of previous work related to multitemporal satellite image classification and the use of phenological data in the interpretation of such data. In section 4 we give a very brief resume of the different types of data used in the project as well as a description of the preprocessing methods applied. The method that we have implemented is described in detail in section 5, results and a discussion of these are given in section 6. Section 7 contains our conclusions.

3 Previous work

The problem of classification of satellite remote sensing data has been studied since the advent of remote sensing satellites. Recently, technological advances in satellite technology has made multispectral and hyperspectral satellite images available, in general this greatly facilitates the task of classifying the underlying scenes. The large number of available satellites has also made fusion of information from different satellites an interesting option when classifying satellite scenes and numerous works report on studies related to this problem, see for instance Solberg et al. (1996) and Benediktsson and Kanellopoulos (1999). Another option arising from the high availability of satellite images is to classify satellite scenes based on temporal sequences of satellite images. It is reasonable to assume that the different classes comprising the scene will have a varying temporal evolution over a growing season. Multitemporal classification methods take advantage of this when basing the classification on an observation of the same scene at several instances in time, see for instance Andres et al. (1994), Lambin and Strahler (1994), Olsson and Eklundh (1994), Running et al. (1995), Lambin (1996), Azzali and Menenti (2000).

Spatial and temporal resolution are, unfortunately, a necessary trade-off in remote sensing. If the given problem requires data with a high spatial resolution, then the temporal resolution will be low and vice versa. The problem considered in this report requires the use of LANDSAT data with a spatial resolution of 30×30 meters. The relatively low temporal resolution of this satellite combined with adverse weather conditions in the study area (clouds obscuring a large part of most LANDSAT scenes) finally results in at most a few (2 to 3) scenes being available per season. This is obviously very little when considering multitemporal classification since the temporal evolution of the classes is poorly resolved by these very few observations.

A possible remedy is the use of *phenological* data derived from sources other than the satellite data. The phenological data would then describe the expected phenological evolution of the different classes in such a way as to facilitate the interpretation of the satellite data.

The study of plant phenology is an old discipline in many countries. Recently, the interest for research in this field has been revived as plant phenology is expected to be a sensitive indicator of global or local climatic changes, see for instance Roetzer et al. (2000), Sparks and Menzel (2002), Scheifinger et al. (2003)van Vliet et al. (2003). In order to consider not only local and regional aspects of plant phenology, but to consider it on a national and global scale, plant phenology studies based on remote sensing have attracted much interest, see for instance Ludeke et al. (1996), Owen et al. (1998), Schwartz and Reed (1999), Chen et al. (2000), Schmidt and Karnieli (2000), Zhang et al. (2001), Schwartz et al. (2002), White et al. (2002), Zhang et al. (2002), Kang et al. (2003), Zhang et al. (2003) and Zhang et al. (2004).

Although many studies (see above) of ground cover classification are based on classifying pixels based on their temporal evolution over a growing season, few authors report on using *phenological models* to support the classification process. The development of different models of plant phenology is extensively reported in the literature, see for instance Chuine et al. (1998), Chuine (2000), van der Meer et al. (2002), Kang et al. (2003), Schaber and Badeck (2003). The use of these models in support of ground cover classification is however

rarely reported in the literature. One exception to this is the work reported in Viovy and Saint (1994). Viovy and his collaborators suggest using a Hidden Markov Model (HMM) of plant phenology and then points out that this model could potentially have value as a tool for classification. The use of HMMs and similar constructs in the study of phenology has been exploited by many authors, see for instance Balzter et al. (1998), Balzter (2000) and Dale et al. (2002). Viovy's suggestion of using HMMs in direct support of classification seems to be unique in the literature. We will discuss Viovy's work in greater detail in section 5.

4 Data

There are two data inputs necessary to this work, satellite images and phenological models. In this section we will discuss both and describe the necessary data preprocessing.

4.1 Ground truth data

The Norwegian Institute of Land Inventory (NIJOS) produces detailed digital vegetation maps for selected regions in Norway. In this study we have used to such maps, those for Venabygd and Sør-Fron.

4.2 Satellite data

Satellite data are used at two levels in this work. The seasonal characteristics are derived from AVHRR and MODIS data. The preprocessing of these data are described in Huseby et al. (2005). The ground cover classification itself is based on LANDSAT data. The application we consider for the methods we have developed is that of ground cover classification in a mountainous region in southern Norway (Venabygdsfjellet, see figure 1). For this region, the growing season is expected to cover May through October. As we have already pointed out, adverse weather conditions in Norway severely reduces the number of successful LANDSAT scenes that can typically be acquired for a particular region. Although the LANDSAT satellites can produce images of the same region every 16 days, the largest number of acceptable images we found for a particular season was three. The available LANDSAT data covering the study area are shown in figure 9 in appendix A.

4.2.1 LANDSAT data preprocessing

We dispose of a total of 17 LANDSAT 5/7 TM/ETM scenes giving acceptable to excellent coverage of our study area. The data are unevenly distributed over the years (from 1987 to 2004) and over the seasons as indicated in figure 9. The scenes stem from various LANDSAT imagery providers, common to them all is that they were acquired in the L1G format as specified by the LANDSAT 7 Science Data Users Handbook, see lan (2005). In this format the LANDSAT data are radiometrically and systematically corrected.

Although the L1G products are georeferenced, the georeferencing applied is not based on the use of ground control points and typically results in residual positional errors on the order of $250m$. This is unacceptable in the current project and we perform a manual georeferencing to improve this. This georeferencing is performed using ERDAS Imagine (version 8.7, by Leica Inc.). Ground control points are selected in the uncorrected images and are matched with points of known position and altitude in a water mask (a water mask based on the M711 series of maps made by The Norwegian Mapping Authority is used in conjunction with a 25m resolution digital elevation model, the DEM25, also by the Norwegian Mapping Authority). A warping and interpolation is then performed using the bilinear interpolation approach, in the resulting image the geographic position of the upper left pixel is known and coincides with a fixed $25m$ grid so as to allow for easy comparison between files. The residual error in these corrected images is on the order of $25m$.

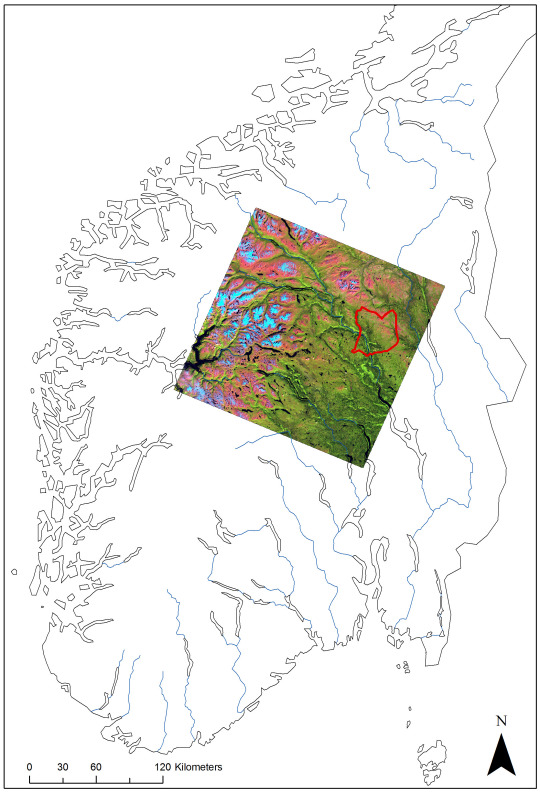


Figure 1: Study area. This figure shows the Norwegian national borders along with one of the LANDSAT (path 199, row 17) images used in this study. The red outline is the border of the Ringebu municipality, most of the study area is included within this border.

In the L1G product, the contents of all the spectral channels are represented as 8 bit digital numbers (range 0 to 255). Before using the images, these digital numbers are scaled back to At Satellite Radiance values. This procedure is described in Chander and Markham (2003) and Ian (2005) for the LANDSAT TM and ETM data respectively.

The final step in the data preprocessing chain is that of calculating the Normalized Difference Vegetation Index. This index is one of the earliest vegetation indexes and has been used in a large variety of applications (see for instance Liang (2004) page 250). It is defined as follows:

$$\frac{\rho_n - \rho_r}{\rho_n + \rho_r}$$

where ρ_n is the at satellite radiance in the near infrared band (LANDSAT TM/ETM band 4) and ρ_r is the at satellite radiance in the red band (LANDSAT TM/ETM band 3).

4.3 Phenological data

Although we combine LANDSAT data from several seasons into one synthetic season with more observations using the methods described in Huseby et al. (2005), the number of

observations per season remains low. Notably, the number of observations will be too low to make high quality ground cover classifications possible directly. In order to improve the classifications, we will incorporate knowledge about the ground cover phenology in the classification process. This will be described in detail in section 5, here we will only briefly discuss the necessary phenological data.

The expected phenological behavior of a standardized vegetation ground cover class when observed in the form of NDVI is illustrated in figure 2 (based on Zhang et al. (2001)). For a vegetation ground cover class with an early development one would expect the curve to be shifted left, similarly a class with a late development should have its corresponding curve shifted right. Different vegetation ground cover types can of course also produce different values of NDVI over the season. The basic idea of the work reported here is that the varying temporal evolutions of different ground cover classes can be used as models in support of the classification process. The problem with this approach is obviously that of defining the models.

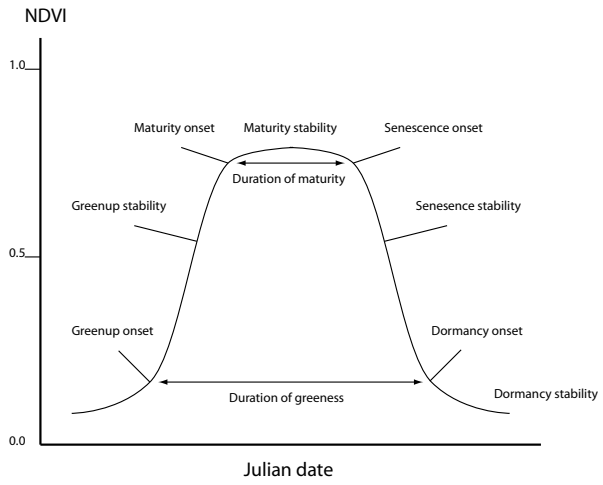


Figure 2: Standard phenological development of a vegetation ground cover class when observed in the form of NDVI.

There is a large literature detailing phenological observations on different plant species from many countries. Typical of these observations is that they pertain to a single species (not vegetation ground cover classes) and they describe very specific discrete events (such as budburst or flowering, not seasonal evolutions). In Norway, such observations are for instance reported in the monumental works by Lauscher and Printz Lauscher et al. (1955), Lauscher et al. (1959), Lauscher et al. (1978), see also Klaveness (1997) and Wielgolaski and Klaveness (1997). The translation of this type of data into models of relevance for the phenology of vegetation ground cover classes is probably not possible. Two options therefore remain: The phenological models can be derived directly from the satellite data or the models can be defined based on expert input.

Both approaches have been extensively explored in this project. The first approach, that of deriving the phenological models from satellite data does not produce satisfactory

results. The main problem is that the LANDSAT NDVI data can obviously not be used for this purpose because the temporal resolution is to low. We must therefore use MODIS or AVHRR data for this purpose. Both satellites can be used to derive NDVI products and the resolution of the final NDVI product i $250m$ and $1000m$ for MODIS and AVHRR respectively. Due to the high granularity of the vegetation cover in the study area, it is impossible to extract class specific phenological profiles based on these data.

The only option left is therefore to derive the phenological models from expert input. This will be discussed in greater detail in section 5.

5 Method

Roughly speaking, our method consists in classifying vegetation ground cover pixels in satellite scenes based on their spectral properties and the temporal evolution of these properties over a season. It is reasonable to assume that different plant species and plant societies will have a different spectral signatures and different evolutions of their 'greenness', that is in chlorophyll content and leaf area index, over a season. Societies with a rapid evolution in spring will thus reach their peak greenness long before societies with a slower evolution. It seems reasonable that this could be exploited in a classification algorithm in order to separate between different plant species and societies.

In this chapter we give a brief overview of the algorithm we have chosen in order to solve the given problem.

5.1 Markov chain based classification

As shown in figure 2 a plant will typically evolve in four phases during a normal season. These phases are: *dormancy, greenup, maturity and senescence*. These phases can be considered as *plant states*, while the plant is in any one of these states it will behave in a particular fashion. The plant states are linked to the NDVI observations one can make of the plant. In the *dormancy state* the observed NDVI values will be stable and low. The *greenup state* is characterized by a rapidly increasing NDVI value. The *maturity state* is characterized by a stable and high NDVI value and, finally, the *senescence phase* will be characterized by rapidly decreasing NDVI values. The NDVI values observed in the different states are in no way *unique* to that state, rather, the NDVI values observed for a given state will only be linked to the underlying state through a probability density for the observable given the state.

These observations have led us to consider a *Hidden Markov Model (HMM)* formulation for the statistical model liking the time of year, plant type and plant state to observed NDVI values. In the following paragraphs we will describe the HMM formalism and link it to our problem. Our presentation is heavily based on Rabiner (1989) and Viovy and Saint (1994).

5.1.1 Basic HMM formalism

In a HMM we observe a system assumed to evolve through a series of different states. Transitions from one state to another happen with certain probabilities. While in a given state the system will produce observables with a certain probability density. We will denote the set of discrete states Q of the internal system by:

$$Q = \{\Phi_1, \Phi_2, \ldots, \Phi_\nu\} \tag{1}$$

where ν is the number of states. Furthermore, the time series of observations, $\bar{X}$ will be denoted by:

$$\bar{X}^T = \{X^1, \ldots, X^T\} \tag{2}$$

where T is the number of elements of the sequence. The unknown state of the process at time t will be denoted E^t, thus $E^t = \Phi_i$ indicates that the process is in state Φ_i at time

t. As we pointed out above, the states are not directly observable, but are related to to observation X^t at times t, $(t = 1, 2, \ldots, T)$ by a probability distribution of measurements:

$$p(X^t \mid E^t = \Phi_i), i = 1, 2, \ldots, \nu \tag{3}$$

For a given time period the model is also described by an array of transition probabilities between each pair of states $p(E^t = \Phi_i \mid E^{t-1} = \Phi_j), i, j = 1, 2, \ldots, \nu$. The probabilities of transition between the different states are obviously strongly dependent upon season, thus the process is not stationary and the matrices of probabilities of transition are time dependent. The final parameters of the model are the initial conditions defined by the probability of being in a given state at the initial time $p(E^1 = \Phi_i), i = 1, 2, \ldots, \nu$.

5.1.2 The HMM formalism and classification

As we stated above, the notion of a *class* from the classification literature becomes the notion of a *system* in the HMM formalism. Traditionally, classification of the vegetation cover observed in a temporal sequence of satellite images is the problem of assigning each pixel in the scene to a *class* based on this pixels spectral properties (or derived properties) and the time evolution of these. In the HMM case, our aim is to assign each pixel to the system that best explains the observed temporal evolution of the pixel. In Rabiner (1989), Rabiner identifies three fundamental problems that can be solved using HMM. In particular, he identifies the following problem: *Given an observation sequence $\bar{X}^T$ and a specific HMM, how do we choose the state sequence that is optimal in some sense, i. e. the state sequence that best explains the observations.* Our classification problem is a simple extension of this problem, we want to optimize the state sequence of each of the HMMs (one per vegetation cover type) and then choose the HMM that best explains the observations. In order to do this we have to solve Rabiner's problem of state sequence choice *for each* HMM and then choose the HMM that provides the state sequence that best explains the observations.

Solutions to this kind of problem are important in many applications and several algorithms are available. They are mainly distinguished by the requirements they impose on data availability. Baum's algorithm Baum et al. (1970) bases decisions concerning time t only on the previous decisions, that is, decisions at time $t = 1, 2, \ldots, t-1$. Viterbi's algorithm (and variants) Forney (1973),Hayes (1975) base the decision at time t on decisions at time $t = 1, 2, \ldots, t-1, t+1, t+2, \ldots, T$. We thus see that Baum's approach is well adapted to *real-time systems* whereas the Viterbi approach is the appropriate choice when all data are available at the moment of decision. For our problem we have chosen to use a method similar to the Viterbi algorithm, see Kitagawa (1987).

5.1.3 HMM formalism applied to our problem

A satellite scene showing vegetation covered ground will almost certainly show several types of plants and plant societies. Each of these societies will, presumably, evolve in a way characteristic for that society. In the HMM formalism, each society represents a *system* with possibly unique hidden states and state transition probabilities. The initial probabilities of being in a given state at $t = 1$ are also unique to each society. If we want

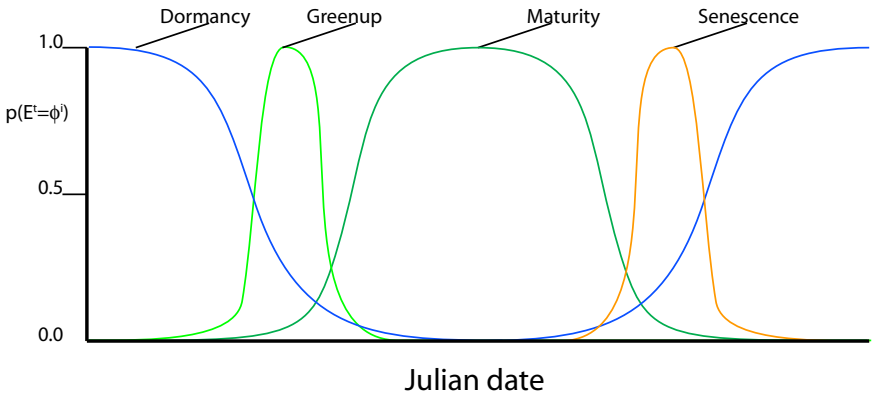


Figure 3: The probability of a given plant society being in a given state as a function of Julian date.

to model for instance three different types of plant societies, we will need three HMM systems, one for each society. We thus see that the notion of a *class* from the traditional classification literature is supplanted with the notion of a *system* in the HMM formalism.

For our problem we will assume that the set of states is common to all the systems describing the various plant societies, they are:

$$Q = \{\Phi_1, \Phi_2, \ldots, \Phi_\nu\} = \{\text{dormancy,greenup,maturity,senescence}\} \tag{4}$$

thus $\nu = 4$. This choice of states is motivated by the typical phenological evolution of a plant society described in figure 2.

The transitions between each pair of these states are described by arrays of transition probabilities as described above. We will assume that conditions are stable over time intervals of D days, system evolution over one year thus requires $365/D$ such matrices (D is typically equal to 7).

Whereas the states are common to all systems, the transition probabilities between states are obviously not common, on the contrary, the properties of the transition matrices along with the probability densities of observables are exactly what distinguishes the different systems from each other. Determining the properties of the state transition matrices should ideally be done *based on data* (the approach chosen in Viovy and Saint (1994)for instance), but as we pointed out in section 4 the scarcity of the available data makes this kind of *training* impossible. In our case we opt for a different approach, we will derive the properties of these matrices from biological and botanical descriptions of the plant societies under study. This approach is based on the observation that knowledge about the probability for a *certain* plant society to be in a certain state given the time of year can be described based on biological and botanical descriptions. Figure 3 shows curves describing the probability of a certain plant society being in one of the four states we consider as a function of the time of year. Assuming that the system *must* be in one of the given states at a given time, the sum of probabilities at a given instance must be equal to one. This obviously imposes very strict conditions on the relationship between these probabilities.

If curves like these can be established for a given society given biological or botanical input, the transition matrices can be derived by the following formula:

$$p(E^t = \Phi_i) = \sum_{j=1}^{\nu} p(E^t = \Phi_i \mid E^{t-1} = \Phi_j) p(E^{t-1} = \Phi_j), i = 1, 2, \ldots, \nu \text{ and } t = 2, 3, \ldots, T \tag{5}$$

Given Norwegian climatic conditions we will assume that all systems are in the dormancy state in the first time interval, that is, the initial probabilities $p(E^1 = \Phi_1) = 1$ and $p(E^1 = \Phi_i) = 0, i = 2, \ldots, 4$.

As we see, this formula describes a set of linear relations between the probabilities of a certain system state vector at time t and the probability of another state vector at time $t - 1$. Solving this set of linear equations makes it possible to find the elements of the transition matrix.

The application of the HMM model to our problem is summarized in figure 4.

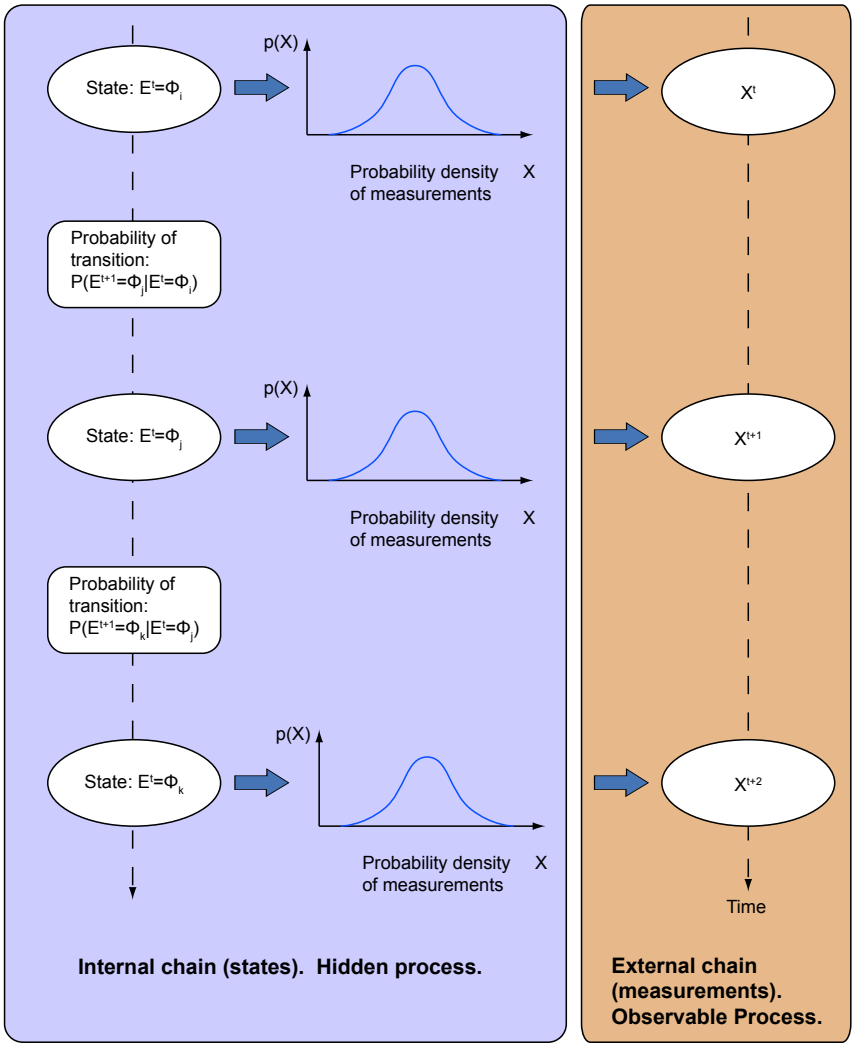


Figure 4: The HMM formalism applied to our problem.

6 Results and Discussion

To test the HMM based classification approach we applied it to a subsequence of the LANDSAT images (7 scenes) covering one summer season (from early June to the middle of October) of southern Norway (the construction of this sequence is the subject of the report by Huseby et. al. Huseby et al. (2005)). The images we used contained six spectral layers each, these were the standard LANDSAT spectra with the exception of the IR spectra. Georeferenced and radiometrically corrected versions of these images were used as input to the subsequent processing. A color composite from the image acquired on the 24th of July is shown in figure 5.

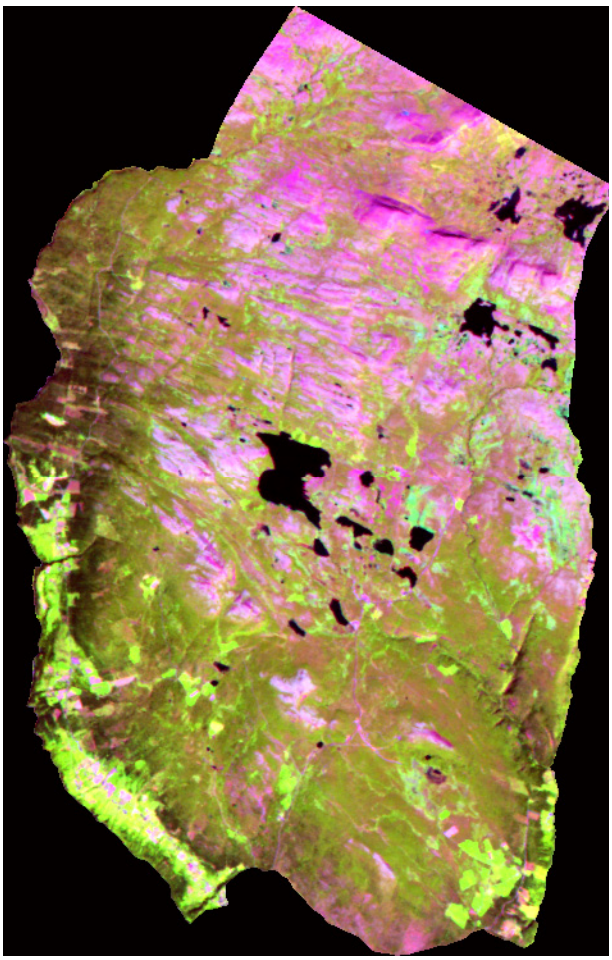


Figure 5: Color composite of the Vena scene based on channels RGB=543.

We also accumulated a stack of all seven images (in chronological order). From these images we chose two test areas, we will refer to these as *Vena* and *Fron*. The two scenes are separated by roughly 10km. Both test areas span a wide range of altitudes (roughly 1000 m). The vegetation ranges from typical Norwegian inland valley vegetation to high mountain vegetation. From both test areas we dispose of detailed digital vegetation maps made by the Norwegian Institute of Land Inventory, (NIJOS), the vegetation map from

the Vena test area is shown in figure 6.

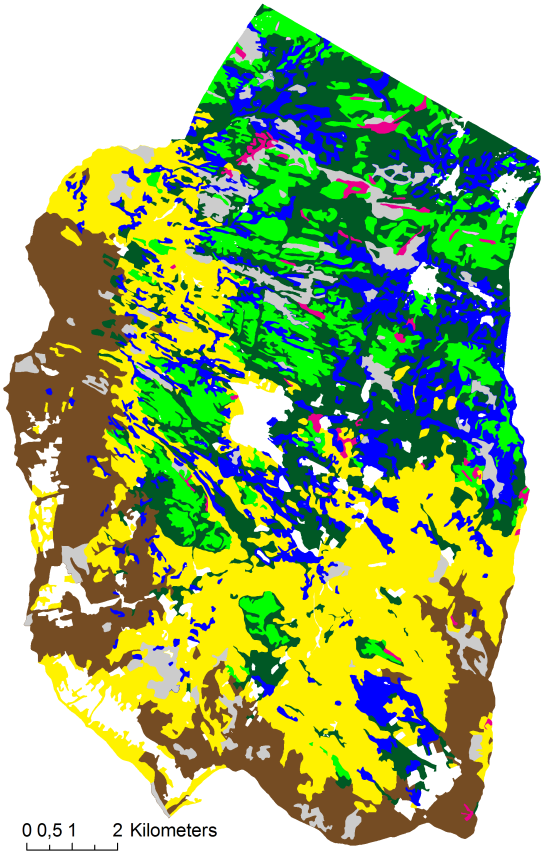


Figure 6: The NIJOS vegetation map from the Vena test area. The color coding is as follows: coniferous is brown, leaf is yellow, rich heather is light green, poor heather is blue, marsh is dark green and rockland is magenta. White areas are either outside the Vena region or areas not interesting for classification (water e.g.). Gray areas are the areas used for training.

We sought to classify the vegetation cover in both scenes into the classes *coniferous* (mainly spruce and pine), *leaf* (mainly mountain birch), *rich heather*, *poor heather*, *marsh* and *rockland*. Using the digital vegetation maps we selected regions representative of each class in the Vena scene. These training regions were then used to establish the necessary class statistics both for each separate image and the stack of all seven images.

A priori probabilities for each class were established (see figure 3) and transition matrices for probabilities of transition between states were established according to equation (5). Classifying the Vena scene using the seven images as input to the HMM method we get an overall classification accuracy of 63.1%. The result is shown in figure 7. Using the HMM approach on the Fron scene, still with training data from the Vena scene, the overall classification accuracy was 62.4%, thus a very slight decrease in relation to the results obtained on the Vena scene.

For comparison both scenes were also classified using the ML approach. Classifying one single image, we obtain the best results using as input the image acquired on the 24th of July. For this image the overall classification accuracy was 58.2% when compared with the digital vegetation maps. The result of an ML classification of the chronological stack of all seven images is shown in figure 8. Using the stack as input the overall classification accuracy was 63.4%, thus using the full stack clearly improves the results. This result is marginally better than the results obtained using the HMM. This is not surprising for several reasons. First of all the ML algorithm was both trained and applied to the Vena region (although the training regions were not included in the regions on which the algorithm was tested). Secondly, the ML approach takes as input the entire covariance matrix for all the bands in the stack of images. Thus all correlations, even between bands stemming from acquisitions on different dates, were known to the ML algorithm. This is obviously not the case for the HMM approach.

In order to test the general validity of the training data acquired in the Vena scene we used the same training data to classify the Fron scene using the ML approach. The overall classification accuracy falls to 56.3% showing that the ML approach is highly sensitive to the validity of the training data. In this case the HMM approach clearly outperforms the ML approach, indicating that the HMM based classification method generalizes better than the ML method. A summary of the classification results is given in table 1

	HMM	ML single	ML stack
Vena	63.1%	58.2%	63.4%
Fron	62.4%	-	56.3%

Table 1: Summary of classification results.

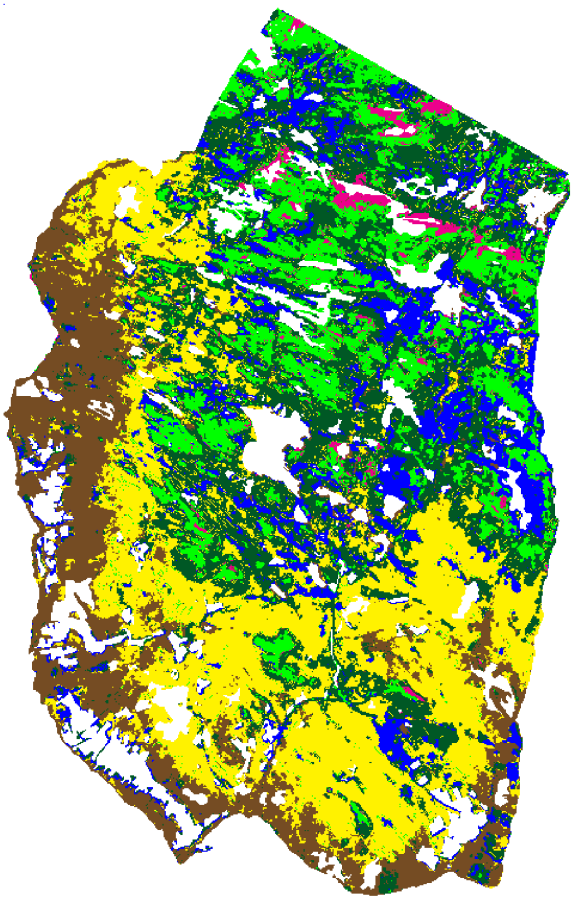


Figure 7: Result of classifying the chronological stack of all images for the Vena scene using the HMM classification algorithm. The color coding is as given in figure 6.

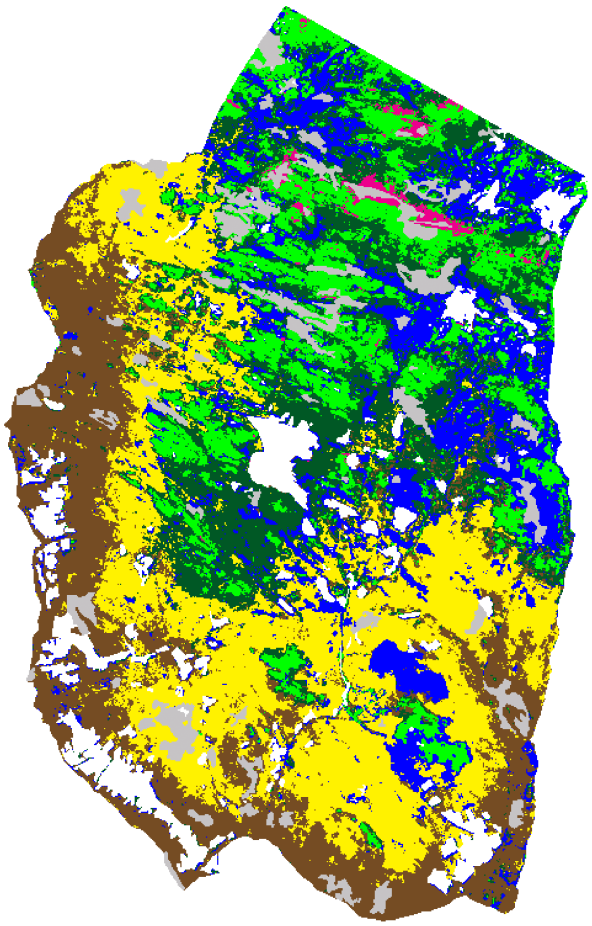


Figure 8: Result of classifying the chronological stack of all images for the Vena scene using a standard ML classification algorithm. The color coding is as given in figure 6.

7 Conclusions

We observe the vegetation ground cover of mountainous regions in a multitemporal sequence of LANDSAT images. In order to classify the sequences, we have developed a new methodology for vegetation ground cover classification incorporating knowledge of phenology into the classification process. The phenological knowledge is represented in the form of a HMM. The classification quality obtained using the HMM approach compares well to that obtained using traditional supervised ML methods. The suggested method is flexible and easily adaptable to other applications.

References

LANDSAT 7 Science Data Users Handbook, 2005. 5, 6

L. Andres, W. A. Salas, and D. Skole. Fourier-Analysis of Multitemporal AVHRR Data Applied to a Land-Cover Classification. *International Journal of Remote Sensing*, 15(5): 1115–1121, 1994. 3

S. Azzali and M. Menenti. Mapping Vegetation-Soil-Climate Complexes in Southern Africa Using Temporal Fourier Analysis of NOAA-AVHRR NDVI Data. *International Journal of Remote Sensing*, 21(5):973–996, 2000. 3

H. Balzter. Markov Chain Models for Vegetation Dynamics. *Ecological Modelling*, 126 (2-3):139–154, 2000. 4

H. Balzter, P. W. Braun, and W. Kohler. Cellular Automata Models for Vegetation Dynamics. *Ecological Modelling*, 107(2-3):113–125, 1998. 4

L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A Maximization Technique Occurring in Statistical Analysis of Probabilistic Functions of Markov Chains. *Annals of Mathematical Statistics*, 41(1):164–171, 1970. 10

J.A. Benediktsson and I. Kanellopoulos. Classification of Multisource and Hyperspectral Data Based on Decision Fusion. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(3):1367–1377, 1999. 3

G. Chander and B. Markham. Revised Landsat-5 TM Radiometric Calibration Procedures and Postcalibration Dynamic Ranges. *IEEE Transactions on Geoscience and Remote Sensing*, 41(11):2674–2677, 2003. 6

X. Q. Chen, Z. J. Tan, M. D. Schwartz, and C. X. Xu. Determining the Growing Season of Land Vegetation on the Basis of Plant Phenology and Satellite Data in Northern China. *International Journal of Biometeorology*, 44(2):97–101, 2000. 3

I. Chuine. A Unified Model for Budburst of Trees. *Journal of Theoretical Biology*, 207 (3):337–347, 2000. 3

I. Chuine, P. Cour, and D. D. Rousseau. Fitting Models Predicting Dates of Flowering of Temperate-Zone Trees Using Simulated Annealing. *Plant Cell and Environment*, 21(5): 455–466, 1998. 3

M. Dale, P. Dale, and T. Edgoose. Using Markov Models to Incorporate Serial Dependence in Studies of Vegetation Change. *Acta Oecologica-International Journal of Ecology*, 23 (4):261–269, 2002. 4

G. D. Forney. The Viterbi Algorithm. *Proceedings of the IEEE*, 61(3):268–278, 1973. 10

J. Hayes. The Viterbi Algorithm Applied to Digital Data Transmission. *IEEE Communications Magazine*, 13(2), 1975. 10

R. B. Huseby, L. Aurdal, D. Vikhamar, L. Eikvil, A. H. S. Solberg, and R. Solberg. Alignment of Growth Seasons from Satellite Data. Technical Report 999, Norsk Regnesentral, 2005. 2, 5, 6, 14

S. Y. Kang, S. W. Running, J. H. Lim, M. S. Zhao, C. R. Park, and R. Loehman. A Regional Phenology Model for Detecting Onset of Greenness in Temperate Mixed Forests, Korea: an Application of MODIS Leaf Area Index. *Remote Sensing of Environment*, 86 (2):232–242, 2003. 3

G. Kitagawa. Non-Gaussian State-Space Modeling of Nonstationary Time Series. *Journal of the American Statistical Association*, 82(400):1032–1041, 1987. 10

D. Klaveness. Botanisk Fenologi pa ville Planter i Norge, de tidligere Norsk Fenologiske Observasjonsseriene og Mulightene for nye Serier. *Blyttia*, 55(2):53–59, 1997. 7

E. F. Lambin. Change Detection at Multiple Temporal Scales: Seasonal and Annual Variations in Landscape Variables. *Photogrammetric Engineering and Remote Sensing*, 62(8):931–938, 1996. 3

E. F. Lambin and A. H. Strahler. Change-Vector Analysis in Multitemporal Space - a Tool to Detect and Categorize Land-Cover Change Processes Using High Temporal-Resolution Satellite Data. *Remote Sensing of Environment*, 48(2):231–244, 1994. 3

A. Lauscher, F. Lauscher, and H. Printz. Die Phaenologie Norwegens. Teil 1. Allgemeine Ubersicht. In *Skrifter*, pages 1–99. Det Norske Vitenskapsakademi, 1955. 7

A. Lauscher, F. Lauscher, and H. Printz. Die Phaenologie Norwegens. Teil 2. Phaenologische Mittelwerte fur 260 orte. In *Skrifter*, pages 1–176. Det Norske Vitenskapsakademiet, 1959. 7

A. Lauscher, F. Lauscher, and H. Printz. Die Phaenologie Norwegens. Teil 3. Tabellen-Karte der Mittelwerte. In *Skrifter*, pages 1–253. Det Norske Vitenskapsakademi, 1978. 7

S. Liang. *Quantitative Remote Sensing of Land Surfaces*. Wiley Series in Remote Sensing. Wiley Interscience, 2004. 6

M. K. B. Ludeke, P. H. Ramge, and G. H. Kohlmaier. The Use of Satellite NDVI Data for the Validation of Global Vegetation Phenology Models: Application to the Frankfurt Biosphere Model. *Ecological Modelling*, 91(1-3):255–270, 1996. 3

L. Olsson and L. Eklundh. Fourier-Series for Analysis of Temporal Sequences of Satellite Sensor Imagery. *International Journal of Remote Sensing*, 15(18):3735–3741, 1994. 3

T. W. Owen, T. N. Carlson, and R. R. Gillies. An assessment of Satellite Remotely-Sensed Land Cover Parameters in Quantitatively Describing the Climatic Effect of Urbanization. *International Journal of Remote Sensing*, 19(9):1663–1681, 1998. 3

L. Rabiner. A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989. 9, 10

T. Roetzer, M. Wittenzeller, H. Haeckel, and J. Nekovar. Phenology in Central Europe - Differences and Trends of Spring Phenophases in Urban and Rural Areas. *International Journal of Biometeorology*, 44(2):60–66, 2000. 3

S. W. Running, T. R. Loveland, L. L. Pierce, R. Nemani, and E. R. Hunt. A Remote-Sensing Based Vegetation Classification Logic for Global Land-Cover Analysis. *Remote Sensing of Environment*, 51(1):39–48, 1995. 3

J. Schaber and F. W. Badeck. Physiology-Based Phenology Models for Forest Tree Species in Germany. *International Journal of Biometeorology*, 47(4):193–201, 2003. 3

H. Scheifinger, A. Menzel, E. Koch, and C. Peter. Trends of Spring Time Frost Events and Phenological Dates in Central Europe. *Theoretical and Applied Climatology*, 74(1-2): 41–51, 2003. 3

H. Schmidt and A. Karnieli. Remote Sensing of the Seasonal Variability of Vegetation in a Semi-Arid Environment. *Journal of Arid Environments*, 45(1):43–59, 2000. 3

M. D. Schwartz and B. C. Reed. Surface Phenology and Satellite Sensor-Derived Onset of Greenness: an Initial Comparison. *International Journal of Remote Sensing*, 20(17): 3451–3457, 1999. 3

M. D. Schwartz, B. C. Reed, and M. A. White. Assessing Satellite-Derived Start-of-Season Measures in the Conterminous USA. *International Journal of Climatology*, 22 (14):1793–1805, 2002. 3

A.H.S. Solberg, T. Taxt, and A.K. Jain. A Markov Random Field Model for Classification of Multisource Satellite Imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 34(1):100–113, 1996. 3

T. H. Sparks and A. Menzel. Observed Changes in Seasons: An Overview. *International Journal of Climatology*, 22(14):1715–1725, 2002. 3

P. J. van der Meer, I. T. M. Jorritsma, and K. Kramer. Assessing Climate Change Effects on Long-Term Forest Development: Adjusting Growth, Phenology, and Seed Production in a Gap Model. *Forest Ecology and Management*, 162(1):39–52, 2002. 3

A. J. H. van Vliet, R. S. de Groot, Y. Bellens, P. Braun, R. Bruegger, E. Bruns, J. Clevers, C. Estreguil, M. Flechsig, F. Jeanneret, M. Maggi, P. Martens, B. Menne, A. Menzel, and T. Sparks. The European Phenology Network. *International Journal of Biometeorology*, 47(4):202–212, 2003. 3

N. Viovy and G. Saint. Hidden Markov Models Applied to Vegetation Dynamics Analysis Using Satellite Remote Sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, 32(4):906–917, 1994. 4, 9, 11

M. A. White, R. R. Nemani, P. E. Thornton, and S. W. Running. Satellite Evidence of Phenological Differences Between Urbanized and Rural Areas of the Eastern United States Deciduous Broadleaf Forest. *Ecosystems*, 5(3):260–273, 2002. 3

F-E. Wielgolaski and D. Klaveness. Norwegian Plant Phenology, a Brief Review of Historical Data and a Comparison of some Mean First Flowering Dates (mFFD) for this and the last Century. *Journal of Biometeorology*, 2:208–213, 1997. 7

X. Y. Zhang, M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, B. C. Reed, and A. Huete. Monitoring Vegetation Phenology using MODIS. *Remote Sensing of Environment*, 84(3):471–475, 2003. 3

X. Y. Zhang, M. A. Friedl, C. B. Schaaf, A. H. Strahler, and A. Schneider. The Footprint of Urban Climates on Vegetation Phenology. *Geophysical Research Letters*, 31(12), 2004. 3

Xiaoyang Zhang, M.A. Friedl, C.B. Schaaf, A.H. Strahler, J.C.F. Hodges, and Feng Gao. Using MODIS Data to Study the Relation Between Climatic Spatial Variability and Vegetation Phenology in Northern High Latitudes. In *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International*, volume 2, pages 1149–1151 vol.2, 2002. 3

Xiaoyang Zhang, J.C.F. Hodges, C.B. Schaaf, M.A. Friedl, A.H. Strahler, and Feng Gao. Global Vegetation Phenology from AVHRR and MODIS Data. In *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*, volume 5, pages 2262–2264 vol.5, 2001. 3, 7

A LANDSAT data coverage

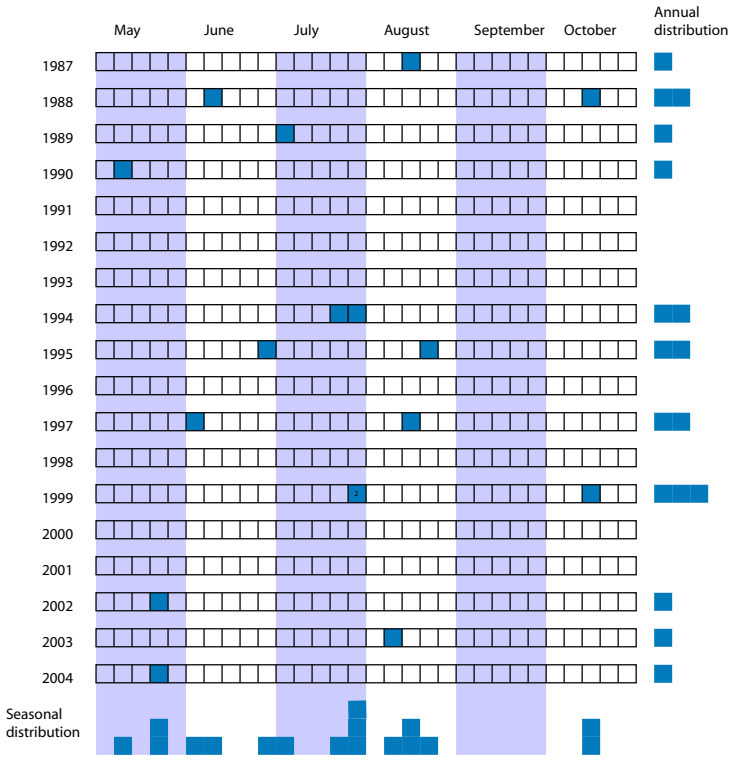


Figure 9: The LANDSAT data coverage of the study area. This figure shows the distribution of LANDSAT scenes that are of acceptable quality for our study area. The distribution is indicated *per year (season)* and *per six-day period of a month*. Notice that no season nor period of a month is covered by more than three images, a result of the weather conditions in Norway.