Discrimination of Saskatchewan prairie ecoregions using multitemporal 10-day composite NDVI data

Xulin Guo, University of Saskatchewan

Abstract: Ecosystems processes include the exchange of water, energy, and greenhouse gases between the soil, vegetation, and the atmosphere. Seasonal characteristics of plants are closely related to characteristics of the annual cycle of weather patterns, therefore, changes in plant phenological events may signal important year-to-year climatic variations or even global environmental change that should be reflected in the differentiation of ecosystems. Monitoring ecosystems that are sensitive to climate change can improve our understanding of the relationships between climate and ecosystem dynamics. This improved understanding is critical for future land-use planning purposes. The objectives of this paper were to characterize the temporal Normalized Difference Vegetation Index (NDVI) values of the four prairie ecoregions and to test the possibility of discriminating these ecoregions using NDVI data. Ten-day composite Advanced Very High Resolution Radiometer (AVHRR) data for Saskatchewan were downloaded from Natural Resources Canada Geogratis website for 1993 to 1998. NDVI channels were subset from each image. Results showed that the NDVI values were unique for different ecoregions as well as different times of year. The NDVI differences among these ecoregions were significant. The results indicated discriminating ecoregions using NDVI data with an overall accuracy of up to 73.5%.

Key words: Saskatchewan, prairie ecosystems, AVHRR, NDVI, MANOVA, discriminant analysis

Introduction

Ecosystems processes include the exchange of water, energy, and greenhouse gases between the soil, vegetation, and the atmosphere. The ability to detect changes in ecosystem processes such as carbon fixation, nutrient cycling, net primary production, and litter decomposition is an important part of defining global biogeochemical cycles. Seasonal characteristics of plants are closely related to characteristics of the annual cycle of weather patterns, therefore changes in plant phenological events may signal important year-to-year climatic variations or even global environmental change. Researchers have been focusing on large-scale changes in terrestrial ecosystems (*e.g.*, Dixon *et al*. 1994; Ojima *et al*. 1994; Lambin and Ehrlich 1997). It is accepted that at the global scale rapid environmental changes are mainly a result of climatic variations and anthropogenic activities. Environmental degradation is also associated with declines in primary productivity that alter biogeochemical exchanges between the earth and atmosphere (Running *et al*. 1994). Monitoring ecosystems that are sensitive to climate change can improve our understanding of the relationships between climate and ecosystem dynamics. This improved understanding is critical for future land-use planning purposes.

Recent advances in remote sensing technology and theory have expanded opportunities to characterize the seasonal and inter-annual dynamics of vegetation communities. Time series analysis of the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR) 1-km multispectral imagery has allowed scientists to examine larger-scale phenological phenomena such as greenup, duration of green period, and onset of senescence (Reed *et al*. 1994), as well as change in seasonally-dependent biophysical variables such as leaf area index (LAI), biomass, and net primary productivity (Roller and Colwell 1986; Gallo and Eidenshink 1988; Achard and Brisco 1990; Teng 1990). Using time-integrated normalized difference vegetation index (NDVI) data, Yang *et al*. (1998) revealed that spatial and temporal variability in growing season precipitation, potential evapotranspiration, and growing degree days are the most important controls on grassland performance and productivity in the central and northern Great Plains.

Temperature increases over the last century within the mixed prairie ecosystem have been among the most dramatic in the world and have resulted in the droughts of the 1930s, 1961, the 1980s, and several others (Wheaton 2000). In southern Saskatchewan, 2001 was one of the driest years in decades, causing severe crop damage (Hayward 2002). Anderson *et al*. (2001) concluded that the temperature in the prairie ecozone of Saskatchewan is expected to increase 3.5^0C to 4.0^0C in the next 50 years. Water availability will decrease because of increasing potential evapotranspiration even with increasing precipitation. Climate change will markedly alter the vegetation regime. Global warming may result in an advance of the northern boundary of C^4 species (Davidson and Csillag 2001) in the mixed prairie ecosystem. From a study conducted by Mitchell and Csillag (2001), precipitation is the primary control factor for vegetation annual productivity in the mixed prairie ecosystem.

However, research into the relationship between vegetation phenology and climate variability has not been fully investigated. This paper set three objectives: 1) to identify the NDVI changes of prairie ecoregions over the growing seasons of six consecutive years; 2) to test the NDVI differences among different ecoregions; and, 3) to investigate the overall accuracy for discriminating ecoregions using time series NDVI data. The paper represents the first part in a series of studies aimed at more fully identifying the relationships between vegetation phenology and climate variables.

Study Area

The study area is southern Saskatchewan. This is bounded by 49^0N latitude in the south, the Boreal forest ecozone in the north, and extends between longitudes 101.5^0 to 110^0W (Fung 1999). The area falls within the prairie ecozone according to the ecological land classification developed in 1991 by the Ecological Stratification Working Group of the federal, provincial and territorial governments. The framework, primarily based on soil, climate and vegetation, comprises three levels of stratification, namely ecozone, ecoregion, and ecodistrict. Ecozone "lies at the top of the ecological hierarchy, and as such, it defines, on a subcontinental scale, the major physiographic features of the country", while ecoregion is defined as "subdivisions of the ecozone, characterized by distinctive climatic zones or regional landforms, and constitute the major bridge between the subcontinental scale ecozones and the more localized ecodistricts" (Acton *et al*. 1998, 3).

The prairie ecozone encompasses four ecoregions extending from the southwest corner of Cypress Upland northward to Mixed Grassland, Moist Mixed Grassland, and Aspen Parkland (Figure 1). The region is dominated by a temperate climate with 1,800 growing degree-days and annual precipitation of less than 300 mm. These dry conditions subside moving northward and eastward to the Aspen Parkland. Prior to European settlement, southern Saskatchewan was covered with natural vegetation, mixed prairie ecosystem. The natural grasslands are fragmented by settlement and agriculture. However, even in the most altered areas, there are pockets of native vegetation which allow visualization of the landscape as it was. Recently, human settlement and, in particular, agricultural development have been the predominant forces in the evolution of the Saskatchewan landscape (Fung 1999).

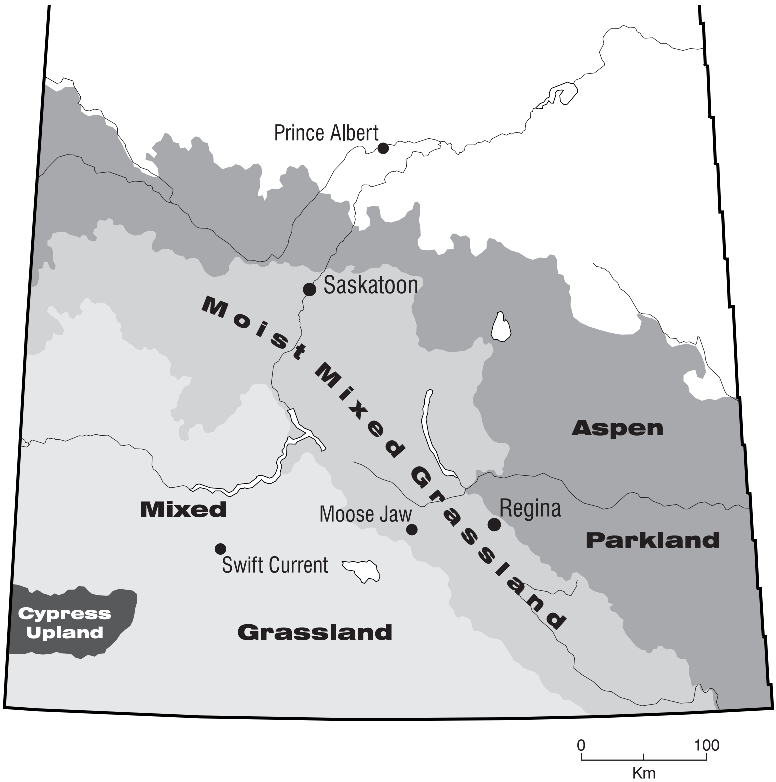

Figure 1: *Study area: the prairie ecozone of Saskatchewan showing its four ecoregions.*

Methodology

Canada-wide 1-km AVHRR 10-day composite maps were derived from NOAA AVHRR data by the Canadian Centre of Remote Sensing (CCRS) for the CCRS Northern Biosphere Observation and Modelling Experiment (NBIOME) project. The original AVHRR data are available each day. The main idea for 10-day composite maps is that after replacing noisy lines using a simple heuristic algorithm, the cloud detection algorithm was used and one image with cloud free or the least cloud cover was selected to represent each 10-day period (Adair *et al*. 2002). Based on the image processing procedure mentioned in the metadata associated with the data downloading, the NDVI layers were computed from the bidirectional reflectance distribution function (BRDF)-corrected surface reflectance for channels 1 and 2. After contaminated pixels were replaced by temporal interpolation of seasonal data, a 5-point smoothing filter was applied to the seasonal NDVI curve. NDVI values from -1 to 1 were scaled to a range between 0 and 20,000.

NDVI, as a vegetation index, is a ratio of the difference between channel 2 (near-infrared wavelength region) and 1 (red wavelength region) to the sum of these two channels. It has the advantages of enhancing vegetation signals, reducing the effects of soil and other non-vegetation features, and standardizing the values. NDVI has been used as a greenness index for vegetation. Thus, the higher the NDVI value, the greater is the greenness of vegetation and the amount of green cover on the ground. This is because green vegetation has higher near-infrared reflectance and lower red reflectance. Bare soil or areas with low vegetation cover have low or negative NDVI values. During the vegetation growing cycle the NDVI value acts as an indicator of the density of chlorophyll on the ground, increases as the vegetation starts to green up, reaches the maximum number at the highest productivity level, and starts to decrease as vegetation becomes senescent. According to Acton *et al*. (1998), ecoregions were classified based on vegetation, climate, and soil, NDVI has the potential ability to signal the vegetation features of different ecoregions and provides valuable information as a remote sensing tool in studying vegetation phenology cycles at a regional scale.

The multitemporal 10-day composite AVHRR data cover the growing seasons (April 11 to October 21) of six years from 1993 to 1998. After downloading the dataset from Geogratis (*http://www.geogratis.gc.ca/*) of Natural Resources Canada (Canadian Centre for Remote Sensing), NDVI layers were subset to the Saskatchewan boundary. Among 2,225 weather stations in Saskatchewan, 141 active weather stations in the prairie ecozone during the 1993 to 1998 period were selected for analysis. Due to the uneven distribution of weather stations, only two were available for Cypress Upland, so an additional six points were selected randomly in this ecoregion to make a total of 147 points for the study area. Pixel values for each date for a 5x5 window surrounding each weather station were extracted from the NDVI image for the four ecoregions using PCI Geomatica V8.1. Points belonging to the same ecoregion were then combined. NDVI values were rescaled back to between -1 and 1.

The median and standard deviation were calculated for each point. Some point locations were observed to be close to water bodies and subsequently were adjusted (shifted away from the water body). This adjustment was made because of the significantly different spectral characteristics of vegetation and water. Clear water has no near-infrared reflectance and higher red reflectance. The purpose of this study was to investigate the vegetation phenology. This process was aided by overlaying the water body coverage for Saskatchewan and visually inspecting each point site. The mean and standard deviation were also calculated for each ecoregion. NDVI curves were plotted for all ecoregions for these six consecutive years. The NDVI layer from April 11 of 1997 was excluded from the dataset due to a georeferencing problem of the downloaded file. Multivariate analysis of variance (MANOVA) was performed to test the NDVI differences among the ecoregions. The best time period(s) to acquire images for separating different ecoregions was/were estimated from the MANOVA test results.

Discriminant analysis was also performed to predict the classification accuracy for different ecoregions. The user and producer classification accuracies were calculated using the methods described by Congalton and Green (1998). Classification accuracy using a *Jack-Knife Cross Validation* (JCV) approach was also performed. This approach was implemented by withholding the spectral data for one point, and building the discriminant functions using the data from the remaining points. The prediction accuracy of the discriminant functions was tested by comparing the predicted ecoregion against the actual ecoregion for the one point that was withheld from the analysis. The process of removing one point from the dataset was repeated 146 times until all points had been withheld and the accuracy of the 146 discriminant functions was tested.

Results

NDVI Variations among Ecoregions and Time of Year:

The NDVI values were plotted along time series for each ecoregion (Figure 2). The graph visually showed that the maximum NDVI values changed from one year to another. The lowest maximum NDVI values were found in 1993 and 1995 and the highest maximum NDVI value occurred in 1996. Among the four ecoregions, Mixed Grassland had the lowest maximum NDVI values throughout the six years. From a geographic perspective, Cypress Upland is at the same latitude with Mixed Grassland but is at a higher elevation, so the vegetation type is more closely related to the Moist Mixed Grassland, which is northward of the Mixed Grassland (Figure 1). The NDVI curves for Cypress Upland and Moist Mixed Grassland were close to each other. Aspen Parkland, an area of mixed

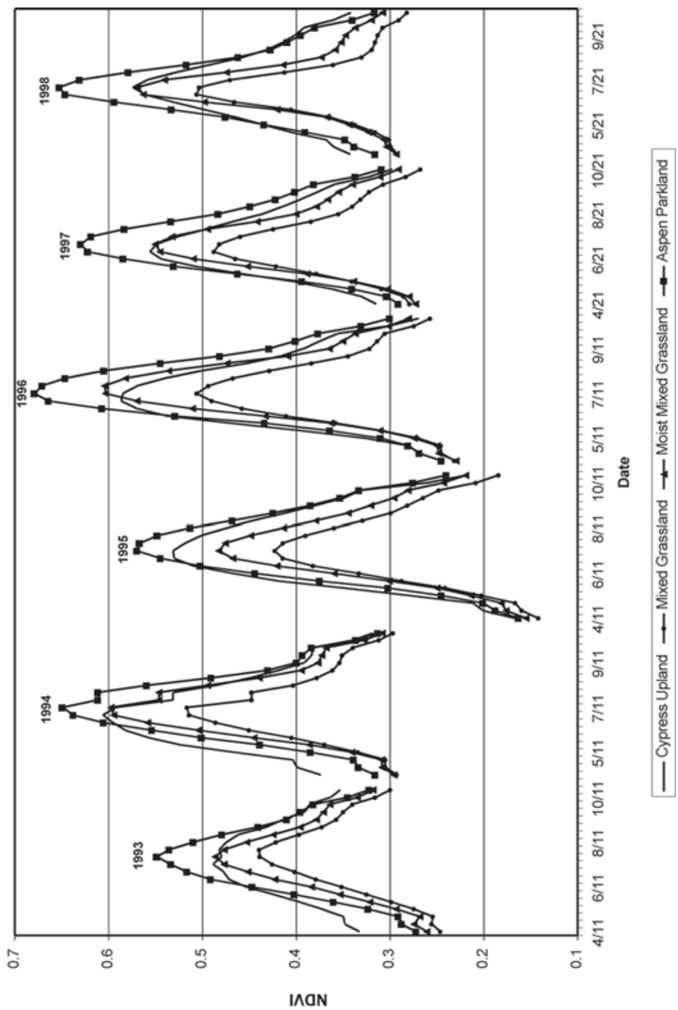

Figure 2: *NDVI values of the four ecoregions for each growing season from 1993 to 1998.*

grasses and trees, is located in the northern part of the region. The spectral near-infrared reflectance of trees is higher than that of grasses. NDVI is the difference between reflectance of near-infrared and red wavelengths over the sum of these two values; therefore, it was not surprising that the NDVI values for Aspen Parkland were the highest. It should be noted that the majority of the prairie ecozone has been converted to cropland. The NDVI values are normally higher in cropland than that in grassland. However, the NDVI value for each pixel is the result of mixed features in that pixel. Based on the lower spatial resolution dataset (1 km x 1 km), each NDVI value might include both cropland and grassland features. Furthermore, the regional climate affects crops in the same way as it affects native grasses.

MANOVA test results:

Overall, the MANOVA test showed the difference among these ecoregions was significant. Subsequently, NDVI differences of each pair of ecoregions were compared using the Tukey Post Hoc test. Results revealed that different pairs of ecoregions had different optimal time periods to be separated. For pairs of Cypress Upland versus Mixed Grassland, Mixed Grassland versus Aspen Parkland, and Moist Mixed Grassland versus Aspen Parkland, the NDVI differences were significant from May to October. To separate Cypress Upland from Moist Mixed Grassland using NDVI data, images from May to middle June were the best. Hot summer (July-August) is the best for separating Cypress Upland and Aspen Parkland. Fall (July-October) images best capture the significant differences between Mixed Grassland and Moist Mixed Grassland. If the Cypress Upland ecoregion was excluded because of its small area and different habitat, the optimal imagery acquisition period to distinguish between the remaining three ecoregions would be hot summer and fall (July-early October). To reduce the length of the table, only the MANOVA test results from 1997 (a normal year) are listed in Table 1.

Discriminant analysis and classification accuracy:

The ability to discriminate among the four ecoregions was tested using NDVI values from these six years. The in-sample validation results indicated that it could discriminate among these ecoregions with 100% accuracy. When validated using the JCV approach, the accuracy level dropped to 73.5%. Table 2 lists classification accuracy results based on the in-sample and JVC approach. From these results, it is clear that the Canonical Discriminant Functions were successful in identifying 108 out

Table 1: *Subset (1997) of MANOVA Tukey Post Hoc test results showing the differences between each pair of ecoregions.*

		pairs of ecoregions					
Month	Day	1 with 2	1 with 3	1 with 4	2 with 3	2 with 4	3 with 4
April	21	**0.008**	**0.001**	0.128	0.808	0.220	**0.020**
May	1	**0.002**	**0**	0.138	0.890	**0.039**	**0.003**
	11	**0.003**	**0.001**	0.465	0.908	**0.001**	**0**
	21	**0.003**	**0.002**	0.955	0.999	**0**	**0**
June	1	**0**	**0.002**	0.993	0.823	**0**	**0**
	11	**0**	**0.001**	0.508	**0.025**	**0**	**0**
	21	**0**	0.131	0.054	**0**	**0**	**0**
July	1	**0**	0.897	**0**	**0**	**0**	**0**
	11	**0**	1	**0**	**0**	**0**	**0**
	21	**0**	0.982	**0**	**0**	**0**	**0**
August	1	**0**	1	**0**	**0**	**0**	**0**
	11	**0**	0.451	**0**	**0**	**0**	**0**
	21	**0**	0.177	**0.027**	**0**	**0**	**0**
September	1	**0**	0.140	0.119	**0.001**	**0**	**0**
	11	**0**	0.184	0.238	**0**	**0**	**0**
	21	**0**	0.311	0.212	**0**	**0**	**0**
October	1	**0**	0.431	0.271	**0**	**0**	**0**
	11	**0.001**	0.667	0.457	**0**	**0**	**0**
	21	**0.012**	0.910	0.626	**0.001**	**0**	**0.006**

Notes: Ecoregions are 1-Cypress Upland; 2-Mixed Grassland; 3-Moist Mixed Grassland; 4-Aspen Parkland
Differences significant at 0.05-level are in bold faces and shaded.

of 147 sites for a total of 73.5% overall accuracy. Both user's accuracy and producer's accuracy showed that the Moist Mixed Grassland was the ecoregion with the lowest discrimination accuracies (Table 2). Figure 3 shows the classification scatter plot for the four ecoregions. The Canonial Discriminant Function 1 adequately separated Mixed Grassland, Moist Mixed Grassland, and Aspen Parkland, while the second function discriminated Cypress Upland from the other ecoregions.

Table 2: *An error matrix showing the classification results according to ecoregion.*

Groups count and %	Predicted Group Membership 1	2	3	4	Total	User's accuracy (%)
Original						
1	**9/*100***	0	0	0	9	*100*
2	0	**39/*100***	0	0	39	*100*
3	0	0	**38/*100***	0	38	*100*
4	0	0	0	**61/*100***	61	*100*
Producer's accuracy (%)	***100***	***100***	***100***	***100***		***100***
Cross-Validated						
1	**7/*77.8***	1/*11.1*	0	1/*11.1*	9	*77.8*
2	0	**31/*79.5***	4/*10.3*	3/*7.7*	39	*79.5*
3	0	5/*13.2*	**24/*63.2***	9/*23.7*	38	*63.2*
4	0	2/*3.3*	13/21.3	**46/*75.4***	61	*75.4*
Producer's accuracy (%)	***87.5***	***79.5***	***58.5***	***78***		***73.5***

Notes: Ecoregions are 1-Cypress Upland; 2-Mixed Grassland; 3-Moist Mixed Grassland; 4-Aspen Parkland
Cells highlighted running diagonally through the table show the number of points correctly and their corrresponding percent accuracy.

Discussion and Conclusion

The study has shown that the NDVI values of vegetation varied among ecoregions and time of year. This may be caused by climate variations but was not tested in this study. The NDVI differences between ecoregions were significant. To capture the maximum NDVI differences among Mixed Grassland, Moist Mixed Grassland, and Aspen Parkland ecoregions, the imagery acquisition date should be between hot summer and fall (July-early October). To spectrally separate Cypress Upland from the other ecoregions an acquisition date in spring (April-early June) or summer (July-August) would be more appropriate. Ecoregions could potentially be classified at 73.5% accuracy overall. Accuracy levels varied depending upon the ecoregions. Moist Mixed Grassland was the one most difficult to discriminate.

There are several limitations in the analysis. First, climate data were not integrated into the analysis, limiting the explanation of the results. Second, the comparison was based on ecoregions, which was too broad a spatial scale to investigate the relationships between plants' temporal characteristics and climate variability. Further analysis is planned that will

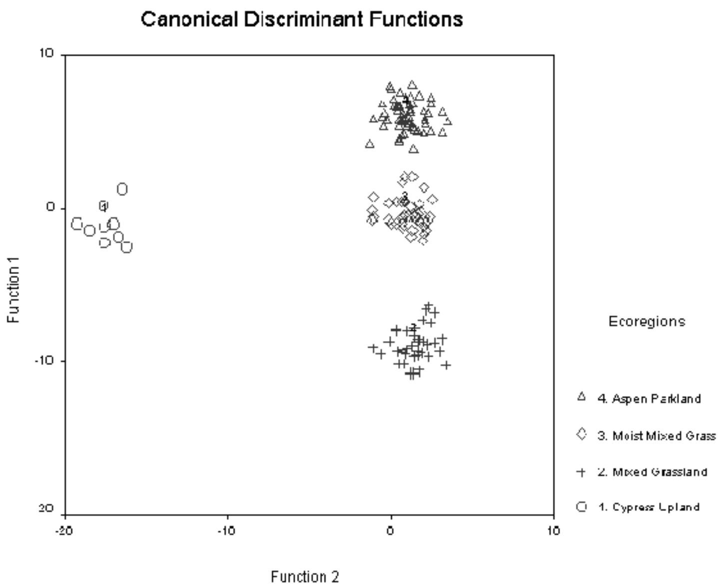

Figure 3: *Canonical discriminant classification scatter plot showing the separability of the four ecoregions using two discriminant functions.*

focus on different land cover types. Third, a six-year time frame is not long enough to conduct climate change analysis, therefore additional NDVI data are needed. These limitations aside, the results provide the basis for conducting further examination of the relationships between vegetation phenology and climate variables. Future research will investigate the relationships between NDVI data that are derived from each weather station and climate variables (e.g., temperature and precipitation). The next step will focus on how climate variables influence vegetation phonological cycles, such as the onset of greenness, the offset of greenness, the maximum NDVI, as well as what climate variable will be most important, and how long the lag will be between vegetation phenology and climate conditions.

Acknowledgements

This study was supported by the start-up funds from the University of Saskatchewan. AVHRR data were downloaded from Natural Resources Canada Geogratis website. I would like to thank Dr. Alain Pietroniro, Elise

Pietroniro, Helton Fan, and Dr. Bill Archibold for their help. Special thanks go to my husband Yunpei Lu for his voluntary efforts in much time consuming work, including AVHRR data downloading, uncompressing, subsetting, and part of the image processing. Finally, I thank the two anonymous reviewers for their valuable comments and suggestions, and Dr. Bernard Thraves for his editorial work on this paper.

References

ACHARD, F. and BRISCO, F. 1990 'Analysis of vegetation seasonal evolution and mapping of forest cover in West Africa with the use of NOAA AVHRR HRPT data' *Photogrammetric Engineering and Remote Sensing* 56:10, 1359-1365

ACTON, D.F., PADBURY, G.A., STUSHNOFF, C.T., GALLAGHER, L., GAUTHIER, D., KELLY, L., RADENBAUGH, T. and THORPE, J. 1998 *The Ecoregions of Saskatchewan* (Regina: Saskatchewan Environment and Resource Management and Canadian Plains Research Center, University of Regina)

ADAIR, M., CIHLAR, J., PARK, B., FEDOSEJEVS, G., ERICKSON, A., KEEPING, R., STANLEY, D. and HURLBURT, P. 2002 'GeoComp-n, an advanced system for generating products from coarse and medium resolution optical satellite data. Part 1: system characterization' *Canadian Journal of Remote Sensing* 28:1, 1-20

ANDERSON, A., JAMES, P., MURPHY, K., ESPIE, R. and GAUTHIER, D. 2001 'The taxa dispersal model: modeling the potential dispersal of organisms in the fragmented prairie ecozone under a future altered climate' in *Proceedings of the GeoSASK2001 Conference*, Regina October 15-17 (Regina: Information Services Corporation)

CONGALTON, R.G. and GREEN, K. 1998 *Assessing the Accuracy of Remote Sensed Data: Principles and Practices* (New York: Lewis Publishers)

DAVIDSON, A. and CSILLAG, F. 2001 'The influence of vegetation index and spatial resolution on a two-date remote sensing-derived relation to C4 species coverage' *Remote Sensing of Environment* 75, 138-151

DIXON, R.K., BROWN, S., HOUGTON, R.A., SOLOMON, A.M., TREXLER, M.C. and WISNIEWSKI, J. 1994 'Carbon pools and flux of global forest ecosystems' *Science* 263, 185-190

FUNG, K. 1999 *Atlas of Saskatchewan* (Saskatoon: University of Saskatchewan) 336

GALLO, K.P. and EIDENSHINK, J.C. 1988 'Differences in visible and near-IR responses and derived vegetation indices for the NOAA-9 and NOAA-10 AVHRRs: A case study' *Photogrammetric Engineering and Remote Sensing* 54:4, 485-490

HAYWARD, K. 2002 'Crop Insurance as a Risk Management Tool for Dryland Agriculture' paper presented at Soils & Crops 2002 conference, Saskatoon, Saskatchewan, February 21-22

LAMBIN, E.F. and EHRLICH, D. 1997 'Land-cover changes in sub-Sahara Africa (1982-1991): application of a change index based on remotely sensed surface temperature and vegetation indices at a continental scale' *Remote Sensing of Environment* 61, 181-200

MITCHELL, S.W. and CSILLAG, F. 2001 'Assessing the stability and uncertainty of predicted vegetation growth under climatic variability: northern mixed grass prairie' *Ecological Modelling* 139, 101-121

OJIMA, D.S., GALVIN, K.A. and TURNER, B.L. 1994 'The global impact of land-use change' *Bioscience* 44, 300-304

REED, B.C., BROWN, J.F., VANDERZEE, D., LOVELAND, T.R., MERCHANT, J.W. and OHLEN, D.O. 1994 'Measuring phenological variability from satellite imagery' *Journal of Vegetation Science* 5, 703-714

ROLLER, N.E. and COLWELL, J.E. 1986 'Coarse-resolution satellite data for ecological surveys' *BioScience* 36:7, 468-475

RUNNING, S.W., LOVELAND, T.R. and PIERCE, L.L. 1994 'A vegetation classification logic based on remote sensing for in global biogeochemical models' *AMBIO* 23, 77-81

TENG, W.L. 1990 'AVHRR monitoring of U.S. crops during the 1988 drought' *Photogrammetric Engineering and Remote Sensing* 56:8, 1143-1146

WHEATON, E.E. 2000 'Canadian Prairie Drought Impacts and Experiences' in Drought *A Global Assessment - Volume 1* ed. D.A Wilhite (London: Routledge) 312-330

YANG, L., WYLIE, B.K., TIESZEN, L.L. and REED, B.C. 1998 'An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the U.S. Northern and Central Great Plains' *Remote Sensing of Environment* 65: 25-37