UNIVERSITY OF CALGARY

Noise Reduction for Time Series of the Normalized Difference Vegetation Index (NDVI): An Empirical Comparison of Selected Techniques

by

Jennifer Nicole Hird

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOGRAPHY
CALGARY, ALBERTA
January, 2008

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Noise Reduction for Time Series of the Normalized Difference Vegetation Index (NDVI): An Empirical Comparison of Selected Techniques" submitted by Jennifer Nicole Hird in partial fulfilment of the requirements of the degree of Master of Science.

Supervisor, Dr. Gregory McDermid, Department of Geography

Dr. Mryka Hall-Beyer, Department of Geography

Dr. Michael Quinn, Faculty of Environmental Design

Date

Abstract

Satellite-derived NDVI time series are fundamental to the remote sensing of vegetation phenology, but their application is hindered by prevalent noise resulting chiefly from varying atmospheric conditions and sun-sensor-surface viewing geometries. A model-based empirical comparison of six selected NDVI time series noise reduction techniques revealed the general superiority of the Double Logistic and Asymmetric Gaussian function-fitting methods over four alternative filtering techniques. However, further analysis demonstrated a strong influence by noise level, strength and bias, and the extraction of phenological variables on technique performance. Users are therefore strongly cautioned to consider both their ultimate objectives and the nature of the noise present in an NDVI data set when selecting an approach to noise reduction, particularly when the derivation of phenological variables is a final goal. This work is crucial to improving current understandings of NDVI time series noise reduction and its role in the remote sensing of vegetation phenology.

Acknowledgements

First, I would like to express my sincere appreciation and gratitude to my supervisor Dr. Greg McDermid, for the continued support, patience and encouragement he has offered me throughout my graduate program. Without his generous dedication of time, knowledge, advice and an unwavering belief in my abilities, this thesis would not have been what it is. Thank you. To the members of the Foothills Facility for Remote Sensing and GIScience, Julia Linke, Adam McLane, David Laskin, Dr. Guillermo Castilla, Chen Gang, and Ryan Powers, I also offer heartfelt thanks for lending me their expertise, assistance and most particularly, their friendship, which have provided me with a very pleasant and most inspirational research environment. Thank you to my thesis proposal and defense committee members, Dr. Darren Bender, Dr. Michael Quinn and Dr. Mryka Hall-Beyer, for their helpful advice and suggestions, and the to latter, especially, for also offering additional friendly guidance and support. To Dr. John Yackel, who first initiated and facilitated my entrance into the graduate program, I would also like to express my gratitude. The present document may not have come into being without this first encouragement. I would like to express my appreciation to Paulina Medori for her great kindness in keeping me, and many others, informed of the relevant deadlines, forms and procedures to be met, filled out and followed. In addition, without the very generous scholarships provided by the Natural Sciences and Engineering Research Council (NSERC) and the Alberta Ingenuity Fund, I would not have been able to pursue my graduate program and achieve what I did. For their invaluable financial support I am extremely grateful to both of these organizations. To those authors who provided me with additional knowledge concerning their published research, Pieter Beck, Dr. João Carreiras, Dr. Jin Chen, Dr. Lars Eklundh, Dr. Per Jönsson, Dr. Mingguo Ma, and Dr. Toshihiro Sakamoto, I offer many thanks; without their help, the current research may not have been such a success. Finally, to my friends and family, who offered their patience, understanding and at times, sympathy, I offer a heartfelt thank you. My husband, in particular, deserves my warmest gratitude for his patience, love and friendship, and for lending me the courage and fortitude to finish what I had begun.

Table of Contents

List of Tables

Appendix D Tables

Appendix E Tables

List of Figures

Appendix A Figures

Appendix C Figures

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
ABC3	Atmospheric, Bidirectional and Contamination Corrections of the Canadian Centre for Remote Sensing
ARMA	Autoregressive Moving Average
ARMD	Autoregressive Moving Median
ARMM	Autoregressive Moving Maximum
AVHRR	Advanced Very High Resolution Radiometer
BISE	Best Index Slope Extraction
BRDF	Bidirectional Reflectance Distribution Function
BRDF-C	Bidirectional Reflectance Distribution Function Composite
CV-MVC	Constrained View Angle Maximum Value Composite
DFT	Discrete Fourier Transform
DMA	Delayed Moving Average
EOS	End of the Growing Season
EVI	Enhanced Vegetation Index
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FFT	Fast Fourier Transform
FMFGBRP	Foothills Model Forest Grizzly Bear Research Program
GPP	Gross Primary Production
HANTS	Harmonic Analysis of NDVI Time Series
HDF-EOS	Hierarchical Data Format - Earth Observing System
IDL	Interactive Data Language
I-NDVI	Integrated NDVI
LAI	Leaf Area Index
LGS	Length of the Growing Season
LP DAAC	Land Processes Distribution Active Archive Center
MODIS	Moderate Resolution Imaging Spectroradiometer
MVC	Maximum Value Composite
MVI	Mean-Value Iteration
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NOAA	National Oceanic and Atmospheric Administration
NPP	Net Primary Production
PSN	Net Photosynthesis
QA	Quality Assurance
QASDS	Quality Assurance Science Data Set
RMSE	Root Mean Square Error
SMN	Seasonal Midpoint NDVI
SOS	Start of the Growing Season
SPOT	Satellite Pour l'Observation de la Terre
SR	Simple Ratio
VI	Vegetation Index

Chapter One: Introduction

The study of recurrent vegetative, biophysical events such as springtime budburst, leaf-out, flowering and autumnal senescence is referred to as *vegetation phenology* (Badeck *et al*. 2004). Agricultural interests in phenology date back thousands of years, but more recent interests in the topic have expanded considerably to include those of biologists, ecologists, climatologists and other researchers (Zhang *et al*. 2004). This interest stems from the recognition that close relationships exist between vegetation phenology and both biotic and abiotic environments. For instance, plant growth and development is affected by air and soil temperatures, solar illumination, moisture and photoperiod (Reed *et al*. 1994). Vegetation phenology is therefore one of the most responsive and easily observable environmental traits that varies in response to climate, rendering it of particular interest in global climate change research (Badeck *et al*. 2004). In addition, animal activity, movement and population dynamics including migratory patterns (Sanz *et al*. 2003), breeding behaviour (Loe *et al*. 2005) and species richness and distribution (Hurlbert and Haskell 2003) can be closely linked to seasonal vegetative growth. Thus, not only does vegetation phenology provide a means of studying the terrestrial effects of climatic variation and change, but also, the effects of land use and land cover change on ecosystem processes and dynamics.

The need for regional- to global-scale data sets in support of contemporary Earth science research initiatives continues to grow. As a source for the synoptic, repeatable views necessary to such research, satellite remote sensing becomes increasingly popular to a wide variety of scientific objectives. Regional- to global-level investigations of vegetation phenology are no exception. The launch of the Landsat Multispectral Scanner in the early 1970s and the subsequent availability of satellite imagery with global coverage initiated the remote sensing of vegetation phenology over larger areas. Beginning with applications in the enhancement of land cover classification and the development of spectrally-based crop calendars (Badhwar 1982, MacDonald and Hall 1980), remotely-sensed phenological information has become increasingly attractive to researchers interested in large-area Earth and atmospheric dynamics. Over the past three decades a variety of multi-spectral sensors have been employed to provide the data

sources necessary for the remote sensing of vegetation phenology, and since the early 1970s this field of research has expanded considerably, particularly with the launch of the Advanced Very High Resolution Radiometer (AVHRR) sensor in the early 1980s, both in terms of the satellite data available and in the research questions involved. Some of these research initiatives include: mapping vegetative transitions in areas of intensive ranching (Dougill *et al.* 1999), detecting crop planting and harvesting dates (Sakamoto *et al*. 2005), linking phenology to other observable biophysical parameters like leaf area index (LAI) or gross primary production (GPP) (Wang *et al*. 2005a, Wang *et al*. 2005b, Turner *et al*. 2003), characterizing vegetation types (Xiao *et al.* 2002), studying vegetation response to ecological change (Pettorelli *et al*. 2005), and linking changes in phenological event timing and duration to climatological shifts (Yang *et al.* 1997).

Spectral vegetation indices (VIs) often form the basis for remote sensing-based phenological investigations (Townshend and Justice 1986, Huete *et al*. 2002). VIs have been closely linked to the chlorophyll content and structural characteristics of green vegetation at the surface (Myneni *et al*. 1995) and are therefore particularly useful for studying vegetation health and development. Multi-temporal data sets of the normalized difference vegetation index (NDVI) are particularly prevalent throughout the literature and continue to be the principal data source for numerous remote sensing-based phenological studies (Pettorelli *et al*. 2005). Its ease of application, long tradition of use and a demonstrably close relationship with photosynthetic activity at the surface (Myneni *et al*. 1995) have contributed to the popularity of the NDVI over the past two decades.

Despite the extensive use of VI time series in the remote sensing of vegetation phenology, the persistent noise inherent in these data sets continues to be of great concern. Varying atmospheric conditions caused by cloud, ozone, dust, and other aerosols, varying sun-sensor-surface angle geometries manifested in bidirectional reflectance, and sensor calibration and co-registration errors all contribute to variations in the data that are unrelated to surficial vegetative phenology or conditions of interest, and are therefore considered noise (Goward *et al*. 1991, Viovy *et* al. 1992, Cihlar *et al*. 1997*,* Huete *et al*. 1999, Roerink *et al*. 2000, Kang *et al*., 2005, Sakamoto *et al*. 2005). This

noise must be minimized before VI data sets can be used either efficiently or effectively for phenological investigations.

Some pre-processing is possible, and in newer data sets is standard procedure before the data is released to the public (Huete *et al*. 1999). However, precise and rigorous atmospheric and geometric corrections are neither practical nor feasible, both because of the size of modern satellite image data sets and because the necessary detailed ancillary information is generally unavailable (Cihlar *et al*. 1997). As an alternative, a multitude of techniques for further reducing noise in multi-temporal VI data sets are found in the literature. These can range from simple temporal compositing and moving average filters, to more complex techniques involving Fourier analysis or wavelet transforms (e.g. van Dijk *et al*. 1987, Menenti *et al*. 1993, Sakamoto *et al*. 2005).

The benefits of temporal compositing for multi-temporal VI data sets became apparent in the 1980s, and continue to be recognized as an important noise reduction method (Holben 1986, Chen *et al*. 2003). Indeed, it is now standard procedure for many publicly-available remotely-sensed VI data sets to be composited before distribution. Despite the application of compositing, however, composited data sets nevertheless still contain remnant noise that requires further consideration, and there is a multitude of strategies for dealing with this remnant noise (Carreiras *et al*. 2003). For instance, van Dijk *et al*. (1987) supported the use of a running median-based filter, Viovy *et al*. (1992) proposed their own best index slope extraction (BISE) algorithm, and multiple authors have applied Fourier transform-based methods for further noise reduction (e.g. Menenti *et al*. 1993, Sellers *et al*. 1994, Moody and Johnson 2001, Wagenseil and Samimi 2006). More recently, function-fitting procedures and least-squares polynomial running window filters have been employed (Jönsson and Eklundh 2002, Chen *et al*. 2004, Beck *et al*. 2006, 2007). The majority of these techniques demonstrates successful noise reduction in the literature, but many techniques are not generally applied elsewhere in the literature than by those who propose them. In fact, little other than the common concern for dealing with noise in VI data sets and recognition of a few of the most well-known studies unites this large body of literature.

1.1 Statement of the Problem

Despite the need for noise reduction in multi-temporal VI data sets for the effective extraction of valuable phenological information, and the amount of literature addressing this issue, no standard or prevailing technique for dealing with this noise presently exists. A multitude of techniques for minimizing noise is found in the literature, and while a few are employed by a number of researchers because of successful application elsewhere, many others are presented and applied only once despite their demonstrated success in those applications. An all-inclusive review of these numerous techniques does not currently exist in the literature, nor does a comprehensive comparison of their relative abilities to minimize noise in an effective manner. With the widespread use of multi-temporal VI data sets in the study of vegetation phenology and the application of these data sets to a variety of Earth and atmospheric science research objectives, a more comprehensive understanding of noise reduction techniques is crucial to the accurate remotely-sensed characterization of vegetation health, growth and development.

1.2 Research Objectives

The current study aims to address not only the present lack of a current, comprehensive review of the literature concerning the reduction of noise in VI data sets, but also the need for a superior understanding of the existing noise reduction techniques. Guided by a series of informal research questions, the goal of this work is to address the following four objectives.

1.2.1 Main Objectives

- To complete a comprehensive review of the current literature concerning the application of NDVI time series in the remote sensing of vegetation, organizing and summarizing the relevant knowledge regarding the acquisition and processing of NDVI time series, particularly noise reduction techniques and the subsequent extraction of phenological variables.

Associated Research Questions:

- What techniques are available? Where and when are they used?
- What is the demonstrated success of these techniques?
- What selection of techniques would be most suitable for empirical comparison in this research?

- To determine which of a selected set of noise reduction techniques provides the most effective method for reducing noise in a multi-temporal NDVI data set covering west-central Alberta, Canada.

Associated Research Questions:

- Is one technique best able to minimize noise while maintaining data integrity, showing a distinct superiority to other selected techniques? If not, why? If so, which one?

- To explore factors influencing the ability of the selected noise reduction techniques to minimize noise and maintain data integrity.

Associated Research Questions:

- Does performance vary with land cover? With the year? With phenological metric? With the level, strength and type of noise? Why or why not?

- To investigate whether the decision to apply noise reduction to NDVI time series is always best.

Associated Research Questions:

- Is it always beneficial to perform noise reduction?
- Under what circumstances is it beneficial, or not beneficial?

- Does the benefit of noise reduction also vary with land cover, noise level, year and metric?

1.3 Thesis Organization

Chapter One introduces the importance of vegetation phenology to a wide variety of research interests, and the use of satellite remote sensing imagery in its study. Four main research objectives and a series of associated research questions to be addressed in the thesis are listed. This is followed by a literature review in Chapter Two which addresses the remote sensing of vegetation phenology, the reduction of noise in the multi-temporal NDVI data sets employed in such research, and the extraction of phenological measures from these data sets which reflect the health and development of surficial vegetation. Chapter Three outlines the methods and procedures followed in performing the present analysis. It includes a description of the study area, the building of a model environment and several evaluation techniques. The results of these analyses are presented in Chapter Four, which itself is organized around the four objectives and associated research questions. A similar organization is also followed in Chapter Five, which presents a discussion of the results presented in Chapter Four and provides answers to the research questions posed above. The final chapter, Chapter Six, concludes with a brief summary of the current work, its contribution to the present literature, and a number of recommendations for the direction of future avenues of research.

Chapter Two: The Remote Sensing of Vegetation Phenology, A Literature Review

The following chapter is divided into several sections, each of which reviews a topic relevant to the current research. Section 2.1 provides a general introduction to the topic of satellite remote sensing in phenological research, highlighting the importance of vegetation phenology, the relevant sensors used in this type of research, and the application of the normalized difference vegetation index (NDVI) to studies of vegetation phenology. A review of Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation index products follows in section 2.2, where data pre-processing and sources of remaining noise and error are reviewed. Because additional noise removal is necessary before the effective application of these data, the third section, 2.3, summarizes the variety of noise reduction strategies presented in the literature for use on multi-temporal NDVI data sets, outlining the wide variety of techniques used by researchers to reduce noise in NDVI time series. Section 2.4 summarizes the use of metrics derived from noise-reduced NDVI time series as a means to characterize surface vegetation phenology through these multi-temporal satellite remote sensing data sets, while the final summary section, 2.5, presents a brief review of the chapter.

2.1 Satellite Remote Sensing in Phenological Research

2.1.1 The Importance of Vegetation Phenology

Vegetation phenology is the study of periodic, vegetative, biophysical events, such as the annual cycles of spring green-up and autumnal senescence (Badeck *et al*. 2004). As a crucial component in all local, regional and global-level ecological dynamics, vegetation phenology has been of considerable interest to numerous branches of research: ecology, biology, agriculture, environmental science, and global change research, among others (Chen *et al*. 2000, Schwartz *et al*. 2002, Turner *et al*. 2003 Ricotta *et al*. 2003, Sakamoto *et al*. 2005, Reeves *et al*. 2006). Firstly, it is a driving force in the lifecycles of innumerable organisms; vegetation phenology plays a critical role in the timing and abundance of food supplies for insects, birds, rodents, ungulates, and many other organisms. For example, the timing of peak caterpillar abundance, which

occurs after spring budburst in deciduous trees, is linked to the timing of reproduction for insectivorous bird species that rely on the abundance in this food supply to fulfill nestling nutritional requirements (Blondel *et al*. 1993), and many ungulate breeding seasons are timed to take advantage of peak plant growth in order to optimize offspring survival (Gaillard *et al*. 2000, Loe *et al*. 2005). Grizzly bear activity and movement are highly dependent upon seasonal vegetation: the spring, summer and autumn shifts in the distribution of high energy food sources influence grizzly presence over a variety of landscapes (Hobson 2005, Munro *et al*. 2005). Vegetation phenology also plays a part in annual cycles of insect outbreaks and disease control. For instance, it has been demonstrated that the spring production of malaria-carrying mosquito larvae are highly correlated with the spring green-up of various plants in wetlands and aquatic environments (Penfound *et al*. 1945), while desert locust breeding and migratory patterns are shown to be influenced by vegetation distribution and growth across the African desert (Despland *et al*. 2004).

In addition to a notable role in biological lifecycles, the very nature of the seasonal shifts in photosynthetically active vegetation renders it an important component in the terrestrial carbon balance. By absorbing CO_2 during the photosynthetic process, green vegetation contributes to local, regional and global CO_2 fluxes and land-atmosphere exchanges (Aurela *et al*. 2001). These exchanges fluctuate throughout the year with annual cyclic shifts in vegetation phenology. Chen *et al*. (1999) showed that while CO_2 was being released into the atmosphere during winter and early spring over an old aspen forest in Saskatchewan, Canada, maximum CO_2 absorption occurred just after leaf emergence, then decreased gradually until the late growing season, when a rapid decrease in CO_2 uptake corresponded with leaf senescence. Such patterns are also reflected globally on much larger scales (Myneni *et al*. 1997). The connection between vegetation phenology and the global distribution and movement of CO_2 is evident, and the link between CO_2 and global climate is well-established (Hall *et al*. 1975).

In light of the obvious need to understand vegetation phenology and its role in ecosystem dynamics, atmospheric CO_2 balance and global response to climatic shifts, its widespread incorporation into a variety of ecological, atmospheric and biological

research objectives is understandable. As a means of providing repeatable, large-area views of the Earth's surface and spectral responses tied to vegetation health and development, satellite-derived remote sensing data sets play an increasingly integral role in phenological research, and are prevalent in the literature.

2.1.2 Satellite Sensors for Vegetative Phenological Research

With the launch of the satellite-borne Landsat Multi-Spectral Scanner in 1972, the applicability of these repeatable, synoptic data sets for the study of vegetation, and particularly vegetation phenology, became apparent (Reed *et al*. 2003). Wiegand *et al*. (1979) used Landsat imagery to relate satellite-derived reflectance to leaf area and biomass over crop and grassland surfaces; Badhwar (1982) derived phenological profiles and the timing of crop emergence with Landsat in order to better classify different crop types; and MacDonald and Hall (1980) employed Landsat in an attempt to forecast crop yields in Asia. While Landsat continues to provide researchers with relatively high spatial resolution data (30 meters) over large study areas for phenological research (e.g. Numata *et al*. 2007), the limited temporal resolution of Landsat (i.e. return period of approximately 16 days) can be problematic in regions with frequent cloud cover (Xiao *et al*. 2002).

With the launch the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR) in 1978 (Jensen 2000), however, the full potential of satellite remote sensing for phenological research began to be realized. This sensor was originally designed for meteorological applications, but the inclusion of the visible red and near-infrared bands – important components in VI calculations – and daily global image acquisition (Jensen 2000), made the AVHRR exceptionally useful for studying vegetation and vegetation phenology. Not only could useful information about vegetation health and distribution be extracted from these data sets, but their high temporal resolution allowed for a study of these characteristics over time. In one of the earliest applications of AVHRR in vegetative research, Norwine and Greegor (1983) employed a multi-date normalized difference vegetation index (NDVI; discussed below) data set to stratify vegetation types across the state of Texas, concluding that AVHRR imagery showed great potential for vegetative mapping applications.

Justice *et al*. (1986) also demonstrated its utility while deriving phenological amplitude and growing season over a variety of vegetation communities in eastern Africa. Reed *et al*. (1994) characterized the annual phenology of various vegetation types by deriving twelve metrics from a multi-temporal AVHRR data set, illustrating the potential of such data for global vegetation monitoring, while Delbart *et al*. (2006) employed AVHRR to study the influence of snow effects on the identification of spring phenology and associated carbon budget calculations. More than 20 years after the first AVHRR sensor was launched, it continues to be an important tool in phenological investigations.

Although not applied as often as the AVHRR, another satellite sensor that has come to the forefront in vegetative research more recently is the VEGETATION sensor onboard the Satellite Pour l'Observation de la Terre (SPOT) satellites 4 and 5, launched in 1998 and 2002, respectively (Lillesand et al. 2008). Unlike the AVHRR sensor, VEGETATION was built with the intention of providing vegetation-focused data sets to researchers by narrowing the visible red and near-infrared bandwidths, thereby eliminating some of the spectral variation unrelated to vegetation biophysical conditions that is present in AVHRR data sets (Jensen 2000). Since the launch of VEGETATION, numerous applications have been published, including mapping forest coverage and type in north-eastern China (Xiao *et al*. 2002), relating satellite and ground observations of leaf area index (LAI) (Wang *et al*. 2005b), estimating the start of season green-up for use in carbon budget calculations (Delbart *et al*. 2005, Delbart *et al*. 2006), and examining vegetation gross primary production (GPP) and fraction of absorbed photosynthetically active radiation (fAPAR) (Wang *et al*. 2004).

Despite the use and applicability of SPOT VEGETATION imagery for phenological research, however, the National Aeronautics and Space Administration's (NASA's) Moderate Resolution Imaging Spectroradiometer (MODIS) sensor, first launched onboard the Terra satellite in 1999, is predicted to ultimately surpass the AVHRR in popularity, for several reasons (Reed *et al*. 2003). First, the spatial and temporal resolutions of particular MODIS products are comparable to those of the extensive AVHRR record, allowing for the continuation of long-term studies using new MODIS products, while additional MODIS products provide similar data sets at a variety

of additional spatial and temporal scales (LP DAAC 2006). As in the case of VEGETATION, MODIS bands are also narrower than those of the AVHRR in order to avoid some of the water absorption-related problems encountered while using AVHRR (Huete *et al*. 1999). Third, the atmospheric corrections, reflectance calibration, and bidirectional reflectance distribution function (BRDF) corrections performed on the MODIS vegetation products before their release greatly increase their effective use and quality (Huete *et al*. 2002). Fourth, international teams of users and remote sensing specialists support these data through continuing validation and monitoring (Huete *et al*. 1999). The final reason for the increasing popularity of MODIS for phenological research is its low cost of acquisition by the public – much of the data is available free of charge from the Land Processes Distribution Active Archive Centres (LP DAACs). Before its popularity exceeds that of the AVHRR, however, a more temporally extensive MODIS archive will need to accumulate in order to support long-term vegetation studies. Nevertheless, in light of its numerous advantages and increasing popularity, MODIS provides the remote sensing data set for the current project, and is discussed in more detail in section 2.2 below.

2.1.3 The Normalized Difference Vegetation Index (NDVI)

Regardless of the sensor used, whether it be Landsat Thematic Mapper, VEGETATION, AVHRR or MODIS, the majority of vegetative research studies make use of these spectral data in the form of VIs – algebraic combinations of the visible red and near-infrared spectral bands, sometimes with addition of other spectral bands, that exploit the variable spectral response of healthy vegetation across these different wavelengths (Huete *et al*. 2002). To summarize, green vegetation shows high absorption of visible red light for the purpose of photosynthesis by chlorophyll-containing components, while near-infrared wavelengths are largely reflected by structural cells in green leaves. This variable spectral response of healthy green vegetation to the red and near-infrared wavelengths is widely known as the ‘red edge’ (Myneni *et al*. 1995). Seasonal changes in the VIs that manipulate this relationship have been shown to closely reflect seasonal changes in the photosynthetic and structural characteristics of surface vegetation; an increase in a VI typically reflects increasing photosynthetic activity. On

the basis of previous research, a direct relationship is assumed to exist between continuous changes in surface vegetation reflectance, detected through a VI, and surface phenological development (Badeck *et al.* 2004).

Jordan's (1969) simple ratio (SR) index was one of the very first VIs to exploit these particular spectral characteristics of healthy vegetation, and is calculated by the equation:

$$SR = \frac{DN_{NIR}}{DN_{red}} \qquad (2\text{-}1)$$

where DN_{NIR} and DN_{red} are the digital numbers of a pixel in the near-infrared and red spectral bands, respectively. The simple act of ratioing these two bands works effectively toward reducing much of the inherent variation in the signal due to calibration, noise, sun angles, and atmospheric effects (Huete *et al*. 1999).

Over the following decades, many other indices have followed the SR, including: the perpendicular vegetation index (Jensen 2000), global environment monitoring index (Myneni *et al*. 1995), the normalized difference vegetation index (NDVI; Deering 1978), the enhanced vegetation index (EVI; Huete *et al*. 1999) and a series of soil-adjusted and atmospherically-adjusted vegetation indices (Huete and Liu 1994, Gilabert *et al*. 2002), designed to minimize atmospheric and brightness-related soil effects (Huete *et al*. 1988). See Jensen (2000) for a review. Of these, the NDVI continues to be the most commonly used and most prevalent in the literature. It is calculated using the equation (Deering 1978):

$$NDVI = \frac{\left(DN_{NIR} - DN_{red}\right)}{\left(DN_{NIR} + DN_{red}\right)} \qquad (2\text{-}2)$$

The values resulting from this calculation range from – 1 to +1, where negative values usually indicate a vegetation absence and approximately 0.80 generally indicates very high vegetative biomass (Myneni *et al*. 1995, Huete *et al* 1999). Time series of NDVI – often employed as per-pixel series of NDVI over time – are particularly useful for phenological research, since they provide a means to capture time-sensitive phenological shifts and vegetation developmental stages. NDVI increases and decreases

with annual cycles of surficial vegetation green-up and senescence, and much of the research involving NDVI incorporates multi-temporal data sets for this reason.

Some of the first large-area applications of NDVI to the study of vegetation phenology were undertaken in the mid 1980s, beginning with a number of studies employing the globally-available AVHRR imagery to examine phenological shifts across the African continent. For instance, in a global examination of AVHRR imagery, Justice *et al*. (1985) demonstrated the value of this daily, coarse-resolution data for phenological research and vegetation monitoring at varying scales and across several continents. They described general phenological patterns over Africa, highlighting known areas of drought, generated phenological profiles for a variety of vegetation types across South America, West Africa and South Asia, compared information extraction from AVHRR data sets at varying spatial resolutions, and illustrated the potential of these data for climate change research (Justice *et al*. 1985). Following this work, Townshend and Justice (1986) plotted NDVI profiles for a number of vegetation types across Africa using AVHRR, and compared the years 1983 and 1984 to highlight inter-annual variability in seasonal dynamics across the continent in a qualitative analysis. Similarly, in a regional study of East African vegetation, Justice *et al*. (1986) used NDVI profiles to study the relationship between bushland phenology and rainfall events, and to estimate growing season length.

Phenological research employing multi-temporal NDVI data sets continued to grow through the 1990's, and included the phenological classification of global vegetation types (Lloyd 1990), modeling phenological profiles and biophysical characteristics of agricultural crops (Fischer 1994a), deriving important phenological events from NDVI time series (Reed *et al*. 1994), studying the relationship between NDVI and surface temperature during various phenological stages in crops (Gupta *et al*. 1997), comparing satellite-derived start of season with modeled phenology (Schwartz and Reed 1999), and the linking of vegetation phenology with long distance biota movement (Gage *et al*. 1999), among many others. Interest continues to expand, with more recent research employing NDVI for phenological investigations involving the study of growing season length derived from NDVI across differing climatic zones of China (Chen *et al*.

2000), the modeling of phenological transition dates for vegetation monitoring (Zhang *et al*. 2003), an examination of intra-seasonal dynamics in Alaskan tundra (Jia *et al*. 2004), the linking of spring phenology to precipitation and temperature in a study of climate change (Zhang *et al*. 2004), the detection of rice crop planting, heading, and harvesting using NDVI (Sakamoto *et al*. 2005), and a study of the effects of inter-seasonal rainfall variation in the dry savannah of Namibia (Wagenseil and Samimi 2006). The long-term, continued popularity of the NDVI for studying vegetation phenology is evident from the literature; the utility of NDVI time series for such studies provides the impetus for the current research.

In spite of its prominent usage in phenological research, NDVI time series are prone to noise and error from a number of sources. While much of this is dealt with as effectively as possible in more recent NDVI data sets, such as those provided by the MODIS sensor (see description below), older data sets – particularly those from the AVHRR – can be significantly affected. Much of the variation in NDVI multi-temporal imagery unrelated to surface vegetation phenology is due to atmospheric effects and to varying sun-senor-surface angle geometries.

Atmospheric scattering and absorption due to cloud cover, water vapour, dust, ozone, and other aerosols present in the atmosphere change daily and generally depress NDVI estimates, potentially causing underestimations of vegetation biomass and photosynthesis as well as short-term fluctuations in NDVI time series unrelated to surface phenology (Gutman 1991, Viovy *et al*. 1992, Huete *et al*. 1999). Holben (1986) explained, for example, that high-frequency variability in water vapour and aerosols, along with cloud cover, cause much of the false lows in NDVI estimations by decreasing near-infrared reflectance. Water vapour is shown to decrease NDVI by approximately 0.02 units, and aerosols by 0.06 to 0.12 units, depending on solar and viewing angles. Goward *et al*. (1991) described some of these effects as they pertain to the AVHRR, stating that satellite-derived NDVI estimates can be up to 30% lower than corresponding ground-based estimates due to atmospheric attenuation. Cloud effects are particularly prevalent in tropical regions, where extended periods of cloudiness and aerosol

contamination from biomass-burning are commonly encountered (Carreiras *et al*. 2003, Kobayashi and Dye 2005).

In terms of sun-sensor-surface geometries, daily variations in sensor view and solar zenith angles, in addition to the anisotropic nature of vegetated surfaces, result in bidirectional reflectance effects, which can, themselves, vary with vegetation type (Gutman 1991). Holben and Fraser (1984) demonstrated this in their simulations of varying viewing and illumination angles across an AVHRR scan line, concluding that off-nadir viewing angles and low sun angle during the winter season often decreased NDVI estimates. However, false highs could also occur in the extreme forward-scanning angles (Gutman 1991). Li and Strahler (1992) explained how variations in vegetative structural characteristics (e.g. shape) and plant density between different vegetation communities resulted in varying bidirectional effects across land covers. These effects add to undesirable high-frequency fluctuations in NDVI time series that are unrelated to surface phenology.

Two additional sources of error that are more specific to the NDVI itself are (i) a tendency for saturation over high biomass areas and (ii) background soil effects (Huete *et al.* 1997). For instance, Holben and Fraser (1984), in studying the use of NDVI data sets for land cover stratification, reported more difficulty in discriminating between high density green-leaf cover types than between low density green-leaf cover types because of this difficulty with saturation. Huete *et al*. (2002) also remarked on the lower range of NDVI values over higher biomass land covers than over lower biomass land covers. On the other hand, Holben (1986) noted problems in assessing vegetation cover over the Sahara desert where vegetation can be sparse, and suggested that background soil mineralogy was a contributing factor due to sparse vegetation. As Huete *et al*. (1988) explained this could have been be due to the addition of near-infrared scattering from soil surfaces to that of green plant surfaces, contributing to unrepresentative (i.e. inaccurate) NDVI estimates. In response to this contamination, a variety of soil-adjusted vegetation indices and enhanced vegetation indices have been developed that attempt to account for these saturation and background effects (Huete *et al*. 1988, Huete *et al*. 1999). However, while these alternative indices may be improvements in these specific respects, the

normalizing nature of the NDVI renders it less sensitive to atmospheric and topographic effects than these alternatives, which themselves require more sensor noise removal than does the NDVI (Huete *et al*. 2002). For longer-term studies employing NDVI data sets, satellite orbital drift due to atmospheric drag and degradation in sensor calibration over time can also be additional causes of error and inaccuracy in NDVI estimates (Gutman 1991). However, these factors are often accounted for in more currently-available data sets (Running *et al*. 1994) and are not of concern in the current study. See Kaufman *et al*. (2000) for a more detailed discussion of these effects.

Despite the issues of noise and error in NDVI data sets, and the presence of numerous alternative vegetation indices designed to reduce some of these effects, the NDVI remains particularly popular for the remote sensing of vegetation phenology. This popularity is no doubt due to its ready availability through a multitude of sensors and at a variety of scales (e.g. 250 meters to 16 kilometres (km) resolution, and daily to monthly composites; Pettorelli *et al*. 2005), its ease of calculation (see Equation 2-1), and its long history of use in vegetative research, beginning with the Landsat sensors. The current MODIS NDVI product is referred to as a 'continuity index' by Huete *et al*. (1999), as it can be used in conjunction with the long-running AVHRR NDVI data set. It is because of this popularity and the resultant widespread use of the NDVI that this index forms the focus of the current research: a contribution to the existing knowledge and use of satellite remote sensing for studies of vegetation phenology can be more effectively made if its focus is this widely-used vegetation index. For this reason, a multi-temporal NDVI data set derived from the Terra MODIS sensor forms the foundation of the current investigation.

2.2 The Moderate Resolution Imaging Spectroradiometer (MODIS)

2.2.1 MODIS and Vegetation Phenology

As part of the Earth Observing System program, NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) was intended as a primary source for a recurrent, global data on the terrestrial biosphere, thereby providing data for various global change research objectives (Running *et al*. 1994). The sensor was first launched aboard the Terra

satellite in December 1999, and then complemented by a second sensor launched aboard the Aqua satellite in May 2002 (NASA 2006*;* Parkinson 2003). With 36 spectral bands recorded at various spatial resolutions (250 m and 500 m) every one to two days, MODIS data forms the basis for a host of terrestrial, oceanic and atmospheric remote sensing data products, each supported by one of several specialized MODIS Science Teams (Justice *et al*. 2002). Of these, the suite of MODIS Land products is of particular interest for phenological research. It includes three vegetation-focused product series: i) MOD13 – vegetation indices (VIs): NDVI and enhanced vegetation index (EVI), ii) MOD15 – leaf area index (LAI) and fraction of absorbed photosynthetically active radiation (fAPAR), and iii) MOD17 – gross primary production (GPP), net primary production (NPP) and net photosynthesis (PSN) (Justice *et al* 2002).

Kang *et al*. (2003) used the MODIS LAI product, integrated with regional meteorological data, to detect the spring onset of greenness across Korean temperate mixed forests and to build a successful, regional phenology model for predicting spatial variability in the timing of this event across the area. Wang *et al*. (2005a) similarly identified three main phenological stages at two temperate deciduous forest sites in Denmark and France: leaf production, constant leaf maturity, and leaf senescence using MODIS LAI data products. In applying MODIS LAI and fAPAR products over semi-arid Senegal, Africa, Fensholt *et al*. (2004) suggested that while both captured vegetative seasonal dynamics as observed on the ground, MODIS LAI more accurately reflected ground-based LAI than fAPAR, which it tended to overestimate. On the other hand, Reeves *et al*. (2006) supported the applicability of the PSN MODIS product for the characterization of seasonal and inter-annual variability in above-ground green biomass across a grassland region of North Dakota, especially as it related to variations in seasonal and annual precipitation. The potential of NPP and GPP MODIS products in monitoring seasonal and yearly trends in vegetative productivity was demonstrated by Turner *et al*. (2006), who showed that no overall bias occurred when these products were compared to ground-based measurements at a variety of ground truth sites across North and South America. In addition, Ahl *et al*. (2006) compared a number of phenological transition dates (e.g. onset of green-up and dormancy) derived from several MODIS

vegetation products (EVI, NDVI, fAPAR, LAI) with ground-based observations, and concluded that these products possess sufficient sensitivity for characterizing surface phenology and vegetative developmental stages.

While the MODIS LAI, GPP, NPP and fAPAR products are all demonstrably useful in studies of vegetation phenology, the two vegetation indices, EVI and NDVI, as explained previously, are particularly applicable to this type of research. For example, Zhang *et al*. (2003 and 2004) employed MODIS EVI time series to map four phenological transition dates –spring green-up, vegetative maturity, autumnal senescence and onset of dormancy – across the north-eastern United States, and across the northern hemisphere between 35^0N and 70^0N, respectively. In both cases, the dependence of the timing of these events and the length of the growing season on latitude and temperature was clearly demonstrated. The warmer temperatures and corresponding longer growing season in urban versus rural areas in North America, Europe and Asia was also made evident (Zhang *et al*. 2004). Xiao *et al*. (2006) showed that maximum photosynthetic activity in seasonally moist tropical evergreen forests across northern South America occurred during the dry season by using MODIS EVI to derive phenological profiles for several evergreen forest sites in the area, while Knight *et al*. (2006) demonstrated the successful application of phenology-based land cover classification techniques across the Albermarle-Pamlico estuary system in Virginia and North Carolina with MODIS NDVI data, revealing the ability of such multi-temporal data sets to allow for sub-pixel information extraction. Sakamoto *et al*. (2005 and 2006) used MODIS EVI time series to derive a variety of phenological stages in rice paddies in Japan and across the Mekong Delta in southern Indochina, respectively. Rice planting, heading, and harvesting dates and rice growing period were all identified, allowing for the regional mapping of crop calendars (Sakamoto *et al*. 2005) and of various cropping systems (Sakamoto *et al*. 2006). While both vegetation index products clearly provide the necessary data sets for studies of vegetation phenology, the following describes the production of MODIS NDVI products in particular, as these provide the multi-temporal data set for the current research.

2.2.2 MODIS NDVI Products

2.2.2.1 Product Generation and Processing

The MODIS NDVI image products are part of the MOD13 data product series, as are the EVI image products, and while the current section focuses on the NDVI in particular, most the procedures described here also apply to the MODIS EVI products. Table 2-1 lists the variety of MOD13 products that are available at a variety of spatial and temporal scales. Each is accompanied by a quality assurance science data set that provides per-pixel quality flags, and by the corresponding reflectance input bands used in the VI calculation (Huete *et al*. 2002). Each contains both NDVI and EVI products at the listed spatial and temporal resolution. The Level 3 VI data sets come in gridded tiles that are approximately 1200 by 1200 km, and are provided in the integerized sinusoidal projection (Huete *et al*. 1999). Each tile's location is identified with a vertical column and horizontal row number that indicates the portion of the globe it covers.

Table 2-1: MOD13 series MODIS vegetation index products available from the Terra and Aqua satellites. Information compiled from LP DAAC (2006).

Terra V004[1] Product	Aqua V004 Product	Terra/Aqua Combined	Characteristics
MOD13A1	MYD13A1	MCD13A1	Vegetation Indices 16-day[2] Level 3 Global 500m
MOD13A2	MYD13A2	MCD13A2	Vegetation Indices 16-day Level 3 Global 1 Km
MOD13A3	MYD13A3	MCD13A3	Vegetation Indices Monthly Level 3 Global 500m
-------	-------	MCD13C1	Vegetation Indices 16-day Level 3 Global 0.05Deg CMG[3]
-------	-------	MCD13C3	Vegetation Indices Monthly Level 3 Global 0.05Deg CMG
MOD13Q1	MYD13Q1	MCD13Q1	Vegetation Indices 16-day[2] Level 3 Global 250m

[1]Data that has undergone systematic, statistically robust level 4 validation

[2]Refers to the composite period

[3]Climate modelling grid

The MOD13 products are produced from MOD09 Level 2 gridded daily surface reflectance data which has been corrected for aerosols, ozone absorption and molecular scattering (Vermote *et al*. 2002), and which are produced from reflectance calibrated Level 1B data (Figure 2-1). The Level 2 gridded data are temporally composited to generate the 16-day MOD13 products, from which the 250 m and 500 m NDVI products are spatially aggregated to produce one-kilometre grids (Huete *et al*. 2002). These are then spatially averaged to produce the climate modeling grid NDVI products. The monthly MOD13 products are temporally averaged from the corresponding 16-day products (Huete *et al*. 2002). In the production of the MOD13 data series, a primary aim of the MODIS science teams is external noise removal through improved atmospheric correction, calibration, cloud and cloud shadow removal and the standardization of sun-sensor-surface geometries (Huete *et al*. 2002).

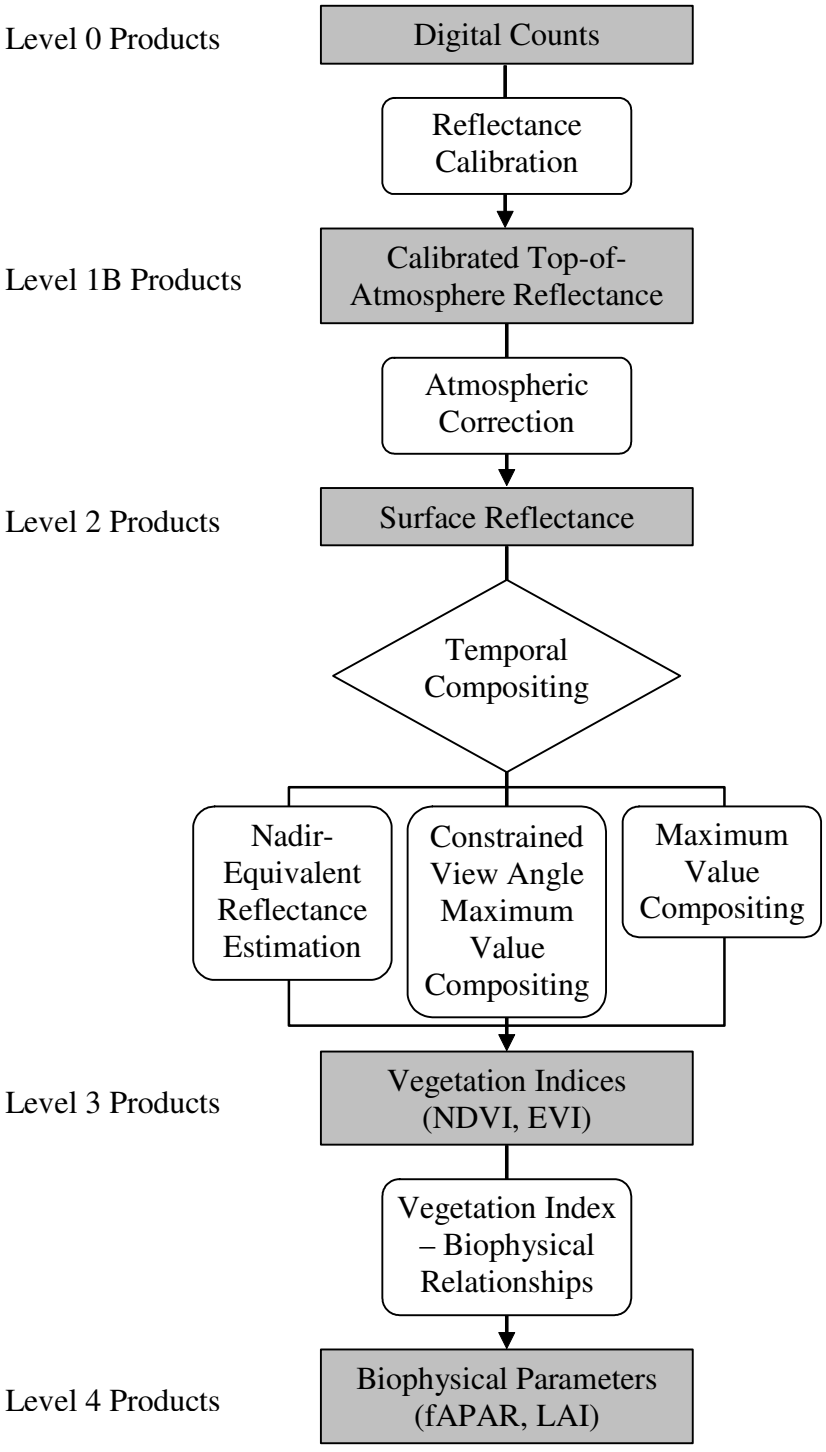

Figure 2-1: Processing steps involved in the production of MOD13 VI products. Adapted from Huete *et al*. (1999).

The MODIS NDVI temporal compositing algorithm functions as a per-pixel operation, relying on multiple observations over the 16-day composite period to generate the final MOD13 products (Huete *et al*. 1999). The number of observations obtained per location within this period may vary anywhere from zero up to 64, due to sensor orbit overlap and multiple observations per day (Huete *et al*. 2002). Once these data are collected, they are filtered into categories based on cloud contamination, view angles and data integrity (Figure 2-2), and only cloud-free, higher-quality data (i.e. pixels with near-nadir view angles and minimal residual aerosol contamination) are used in the subsequent compositing (van Leeuwen *et al*. 1999). Because the whiskbroom nature of the MODIS sensor produces scan angle-related distortion (i.e. pixel size increases by as much as a factor of four; Huete *et al*. 2002), observations with nadir-view scan angles provide the least distortion and are therefore the most desirable for temporal compositing. Out of a possible 64 observations, the number of acceptable pixels covering a 16-day period is generally less than 10 and frequently less than five, most often because of prevalent cloud cover (Huete *et al*. 2002).

Once the data have been filtered and the high quality, acceptable pixels chosen, the data undergo procedures in one of the three components of the MODIS compositing algorithm: i) a bidirectional reflectance distribution function composite (BRDF-C) algorithm, based on Walthall *et al*.'s (1985) BRDF model, ii) a constrained view angle maximum value composite (CV-MVC) algorithm, or iii) a simple maximum value composite (MVC) algorithm (van Leeuwen *et al*. 1999). The choice of which component is used depends on the number of acceptable observations available for a particular 16-day period (Figure 2-2), but in each case the goal is to select the highest quality data to represent that pixel for the composite period. In the BRDF-C component a minimum of five acceptable quality pixels are used to interpolate nadir-equivalent band reflectance values, while for two to five acceptable pixels (Figure 2-2) the CV-MVC method chooses the two observations with view angles closest to nadir, and of those two, chooses the observation that produces the highest NDVI to represent that composite period (van Leeuwen *et al*. 1999). If only one acceptable pixel is available, that pixel is automatically chosen to represent the composite period, while a lack of any acceptable pixels for the

composite period results in the MVC method which simply chooses the pixel with the highest NDVI value calculated for the composite period (van Leeuwen *et al*. 1999).

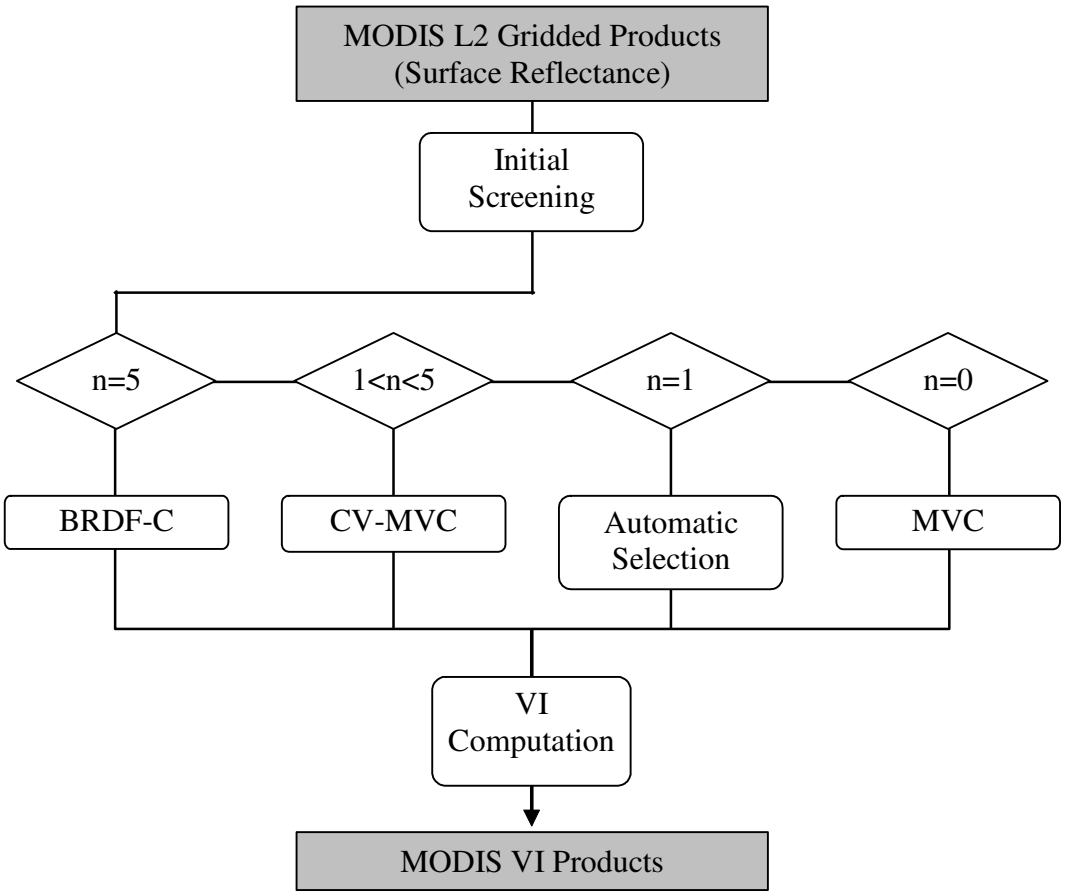

Figure 2-2: Flowchart demonstrating the decision-making process of the MODIS temporal compositing algorithm, where n= the number of acceptable observations after initial screening. Adapted from Huete *et al*. (2002).

As mentioned above, a set of additional scientific data are also provided along with the NDVI products, including quality assurance (QA) information, red, near-infrared, middle-infrared and blue band reflectance (i.e. bands 1, 2, 6 and 3 respectively) and view zenith, solar zenith and relative azimuthal angles of the pixels selected in the compositing process (Huete *et al*. 2002). All MODIS Land products are validated using similar but independent products from in situ, airborne and space-borne sources (Morisette *et al*. 2002). For example, the MODIS Land Quick Airborne Looks aircraft system provides relatively low-cost, versatile source for validation products, while the international network of Earth Observing System Land Validation Core Sites provide a variety of land-based ground truth data (Morisette *et al*. 2002). The results of ongoing

validation are provided both through published literature (e.g. Morisette *et al*. 2002, Gao *et al*. 2003) and quality assurance data included with MODIS Land data sets.

2.2.2.2 Uncertainty and Error in MODIS NDVI Products

Table 2-2 lists the various potential sources of error in the MODIS NDVI products, sorted by processing step, as they are listed by Huete *et al*. (1999). One can see that each step in the process has been carefully considered in terms of these sources of its potential uncertainties. For instance, error related to the reflectance calibration processing steps is found to vary in NDVI products with reflectance (e.g. pixel brightness) and atmospheric conditions, with the largest uncertainties occurring during periods of higher atmospheric turbidity and over dark or sparsely vegetated surfaces (Miura *et al*. 2000). This error is also the result of band to band co-registration and spectral band shifts (Table 2-2). Due to the coarser spatial resolution and aerosol optical thickness heterogeneity at the 250 m, 500 m and 1 km MODIS grid product level, uncertainty can be introduced during aerosol correction on MODIS data because these data rely on climatology grid-based meteorological data (Huete *et al*. 2002). As indicated in Table 2-2, leaf angle, density, and other structural properties can also be a source of error and uncertainty in MODIS NDVI products, as these vary between biomes, ecosystems, vegetation communities, and across time.

Of the various sources of uncertainty, atmospheric sources of error, namely cloud contamination and aerosol-related effects, and sun-sensor-surface geometry-related error, expressed as the bidirectional reflectance distribution function (BRDF), are the most influential on MODIS VI product quality and integrity (Huete *et al*. 1999). While considerable processing and validation efforts are aimed at reducing the effects of these on final MOD13 products, such sources of uncertainty in particular result in variable data quality both spatially and temporally, and quality assurance information must therefore be taken into account when using MODIS-derived NDVI time series for phenological research (Gao *et al*. 2003). Of particular importance for such time series is an understanding of the sources of error and uncertainty that are introduced during the MODIS temporal compositing procedure (Table 2-2). As described in the previous

section, there are three main components to the MODIS VI compositing algorithm, and each with its own advantages and potential sources of error.

Table 2-2: Potential sources of uncertainty and error in MODIS VI products. Adapted from Huete *et al*. (2002).

Proccessing Step[1]	Potential Sources of Uncertainty
Reflectance Calibration	Calibration uncertainty Spectral band shift Band-to-band Coregistration
Atmospheric Correction	Aerosols Residual/partial cloud cover Water Vapor Ozone Pressure
Temporal Compositing	Geolocation Nadir reflectance estimation: i) View/solar geometries ii) Residual/partial cloud cover iii) Shadows iv) Angular effects on CV-MVC
VI-Biophysical Relationships	Canopy backgrounds Canopy structural properties: i) Biome types ii) % green cover iii) Leaf area index iv) Leaf angle distribution

[1] Refer to Figure 2-1 Adapted from Huete *et al* . (2002)

All vegetation indices (VIs) are influenced by the anisotropic nature of vegetated land cover surfaces (Cihlar *et al*. 1994), and the BRDF continues to be of notable concern and interest to the remote sensing community. Bidirectional reflectance is of particular concern to MODIS data products because of the atmospheric correction these data undergo. That is, while atmospheric correction can greatly reduce the effects of atmospheric contamination, it also enhances the effects of land cover anisotropy and varying sun and view angles (van Leeuwen *et al*. 1999). For this reason Walthall *et al*.'s (1985) BRDF model forms the basis for the first component of the MODIS compositing algorithm, aimed at standardizing the MODIS NDVI to nadir view angles and

representative or constant solar angles (van Leeuwen *et al*. 1999). Unfortunately, several factors contribute to the presence of errors in BRD-composited data. First, the MODIS whiskbroom sensor produces varying pixel sizes that are enlarged by a factor of four (Huete *et al*. 2002). This can introduce non-phenology related variation because different sized pixels may have different BRDF characteristics (van Leeuwen *et al*. 1997). Secondly, because this method requires the interpolation of nadir view reflectance from several high quality, cloud-free observations it is very reliant on the accuracy of the MODIS cloud-mask and atmospheric corrections (Huete *et al*. 2002), and errors in either will produce inaccurate composited reflectance and VI values. For instance, Gao *et al*. (2003) found some inconsistencies between MODIS quality assurance labels and observed conditions (i.e. a 'cloudy' designation on a clear day) which were accounted for as an artefact of the cloud mask algorithm. Gao *et al*. (2003) explained that the cloud mask algorithm has the potential to falsely classify bright desert surfaces as cloud, which can lead to errors in the BRDF-C composites. Thirdly and finally, a minimum of five acceptable observations is needed for the BRDF-C algorithm (Huete *et al*. 1999), a number that is not often reached in areas of prevalent cloud cover. This last factor constricts the use of the BRDF model to dry periods and areas of low cloud cover, which significantly reduces its usefulness considering vegetation is most active in wet, rainy conditions (Huete *et al*. 2002). This has ultimately led to the temporary cancellation of BRDF-C use in the MODIS compositing process (Huete *et al*. 2002).

With the BRDF-C algorithm not presently in use, the CV-MVC and MVC algorithms provide the basis for MODIS product temporal compositing. The MVC was primarily designed for minimizing the significant atmospheric effects that are found in AVHRR-derived NDVI, and appears to work well over near-Lambertian surfaces where atmospheric contamination and path length are the principle causes of temporal pixel variation not related to vegetative phenological changes (Holben 1986). However, significant error can be introduced by the BRDF and varying sensor and sun angles, which is only enhanced by the atmospheric correction of MODIS Land data sets (van Leeuwen *et al*. 1999). That is, the MVC can increase the bias toward forward scatter view angles in MODIS products which may not always be cloud-free and which can

often produce higher VI values than nadir-view observations (i.e. inaccurate NDVI estimations; Huete *et al*. 2002). In addition, because pixel size can vary by a factor of four, a bias toward larger view angles results in larger pixels and lower spatial resolutions for the resulting composites. Ideally, the selection of nadir views would produce the most accurate and reliable NDVI computations at the highest spatial resolution. However, this is not guaranteed by the use of the MVC method.

The CV-MVC was developed in order to deal with some of the angular variation and bias that can be introduced by the MVC, particularly in MODIS data sets. By constraining the view angles, the likelihood of angular variation through selecting off-nadir view observations during the compositing process is reduced. However, it should be noted that while van Leeuwen *et al*. (1999) supported the performance of the complete MODIS compositing algorithm as an improvement over the simple MVC method, only slight differences between the results of the CV-MVC and MVC were found when applied to the same data set. Significant improvement was shown by the use of the BRDF-C algorithm over the MVC, but as mentioned before, the BRDF-C was only applicable to limited regions of the world because of its reliance on a minimum of five observations of acceptable quality. Gao *et al*. (2003) also supported the CV-MVC method as superior to the MVC, but did note the presence of some off-nadir observations and residual cloud and cloud shadow contamination. It was also observed that because of the multi-component nature of the MODIS compositing algorithm, more spatial variation due to differing quality levels may be introduced (i.e. different pixels composited using different methods can lead to differing quality; Gao *et al*. 2003). This can be especially true where only a single pixel of acceptable quality is available for a particular composite period and is chosen as the representative for that period, regardless of view angle or calculated NDVI – this could result in highly variable spatial and temporal quality within a MODIS NDVI time series (Huete *et al*. 2002).

Uncertainty is inevitable in any remotely-sensed data set, and while the MODIS Land products have been corrected for the effects of atmospheric contamination and varying sun-sensor-surface geometries, it is recognized that the accuracy of the data must be accounted for through quality assurance and validation activities (Huete *et al*. 1999).

Indeed, Huete *et al*. (2002) expressed the inevitability of varying view angles and residual atmospheric and cloud effects as being aggregated in MODIS composite products from atmospheric correction and the compositing routines. Consequently, QA or quality control flags and information on errors, artefacts and accuracy as well as view and solar angles and snow and cloud presence are provided with MODIS Land products (Justice *et al*. 2002). These are intended to provide the user with an appropriate context within which to consider the reliability, credibility, quality and limitations of the data for their particular research purposes (Roy *et al*. 2002). See Roy *et al*. (2002) for a more detailed description of MODIS product quality assurance.

Ideally only the highest quality pixels with little atmospheric or cloud contamination and nadir views would be used in phenological information extraction – Gao *et al*. (2003) noted that more distinct phenological time series are provided by high quality NDVI observations than by 'acceptable' observations, which lead to increased data heterogeneity and error in inter-annual comparisons – but limiting studies to these data would generally result in large data gaps, restricting their usefulness. It is thus necessary for the available QA information to be considered and MODIS NDVI multi-temporal data sets, including observations of variable quality, to be homogenized in order for these to become sufficiently useful in phenological research.

Because these varying sources of error and uncertainty typically result in high-frequency fluctuations in times series of NDVI that do not reflect the lower frequency vegetative changes ideally observed in such time series (van Dijk *et al*. 1987), this error is generally referred to as noise. Techniques for dealing with and reducing this noise in NDVI time series while retaining the temporal signal of interest (i.e. that related to surface vegetation and its growth and development) are the subject of the following section.

2.3 Noise Reduction in NDVI Time Series

As described previously, multi-temporal NDVI data sets, often employed as per-pixel NDVI time series, are commonly used in the study of vegetation phenology through remote sensing. Of prevalent concern in such applications is the considerable temporal variation resulting from varying atmospheric and bidirectional effects unrelated to

phenological changes, which can exist in these data and limit the effective extraction of phenological information from NDVI time series. The noise resulting from such effects must be minimized before these time series can be effectively applied to the study of vegetation phenology. As a result, a fairly substantial amount of research into the reduction of such noise is found in the literature, and a wide variety of noise reduction techniques have been applied to multi-temporal NDVI data sets. As Cihlar *et al*. (1997) explained, precise, rigorous atmospheric and geometric corrections of these data sets on a per-pixel basis are not practical, nor are they feasible due to a lack of necessary detailed ancillary information. As a result, the multitude of methods for dealing with noise often involves more general time series noise reduction strategies. These generally rely on a few common assumptions.

The two most prevalent assumptions that form the basis of NDVI time series noise reduction strategies are that: i) low-frequency seasonal variation in NDVI is directly related to changes in surface vegetation phenology, and ii) high frequency (e.g. daily or sometimes weekly) fluctuations in NDVI are not generally related to phenological shifts, but rather, reflect short-term variations in atmospheric conditions and sun-sensor-surface view angle geometries (van Dijk *et al*. 1987). Another less common but still popular assumption that forms the basis of various noise reduction techniques is the recognition that much of the noise in NDVI time series is found in the form of spurious lows. That is, because the presence of water vapour and aerosols in the atmosphere lowers the detection of near-infrared reflectance by the satellite, these tend to decrease NDVI values (Holben 1986). As a result, it is often assumed that higher NDVI observations are generally the most accurate, and noise reduction techniques thus aim to preserve the upper envelope of NDVI values in a times series.

One of the most popular and simplest methods for reducing noise in NDVI time series is through multi-temporal compositing (Goward *et al*. 1991, Chen *et al*. 2003, Kobayashi and Dye 2005). Various methods of compositing exist, but they each involve selecting the most desirable date or NDVI value for any given period of time in a time series, based on some specified criterion for that pre-defined period of time, and generally rely on all three of the assumptions listed above (van Dijk *et al*. 1987, Swets *et*

al. 1999, Carreiras *et al*. 2003). Ideally, in this way, each pixel is represented by the most accurate NDVI observation from within the composite period, and daily or sometimes weekly fluctuation resulting from cloud and aerosol contamination and bidirectional reflectance effects, hence unrelated to surface phenology, are removed, while the low frequency phenological shifts occurring at the surface are retained in the signal. The length of the composite period generally ranges from weekly to monthly. While it is recognized that longer composite periods result in better noise reduction, they also limit the extraction of phenological information from time series because of a decreased sensitivity to the timing of phenological transitions with a coarser temporal resolution (Viovy *et al*. 1992, Chen *et al*. 2003).

Compositing is most often done using the maximum value compositing (MVC) method described in Holben (1986), whereby the highest NDVI value for a given composite period is selected to represent that period, as this is assumed to reflect the most cloud-free, aerosol-free, near-nadir view observation. The reduction in atmospheric and bidirectional effects in NDVI time series resulting from the use of the MVC method has been clearly demonstrated (Cihlar, *et al*. 1997, Maxwell *et al*. 2002, Chen *et al*. 2003). A variety of alternative temporal compositing strategies are also found in the literature, such as the constrained view angle MVC (CV-MVC) used in compositing MODIS vegetation index products (Huete *et al*. 2002) and Taddei's (1997) maximum value interpolated compositing method (also see Cabral *et al*. (2003) and Carreiras *et al*. (2003) for additional methods). However, because of its simplicity and popularity, the MVC method is the means through which most NDVI data sets are provided to users as pre-composited data sets (e.g. SPOT VEGETATAION, AVHRR and MODIS NDVI products); as discussed above, MODIS NDVI data sets are composited using a combination of CV-MVC and MVC procedures. Nevertheless it should be recognized that despite the smoothing effects of temporal compositing, residual atmospheric and bidirectional effects often still exist in an NDVI time series, particularly in regions of prevalent cloud cover or at higher latitudes where sun and sensor view angles can be large (Carreiras *et al*. 2003). For this reason, further noise reduction is often performed on composited NDVI time series, and this forms the focus of the following summary. It

can be assumed that the noise reduction techniques described below all pertain to composited NDVI data sets unless otherwise noted.

The majority of NDVI time series noise reduction strategies can be placed into one of two categories: function-fitting or filtering. The first type involve the construction of a new NDVI time series that is fitted to the overall trends in the original profile, that may pass through some, but often few, of the original data points, and that can be described using a function. The second type, on the other hand, are not described by a function, but rather rely on properties of the original data points and a set of rules to replace each data value by a linear combination of a set of nearby values (Jönsson and Eklundh 2006). Filters aim to minimize the high-frequency noise while retaining the desired signal in the times series. Filtering techniques often incorporate more of the original data in their result, but these results are not generally as smooth as those produced through function-fitting (van Dijk *et al*. 1987). On the other hand, function-fitting techniques can produce spurious oscillations that distort the time series, while filters do not (Chen *et al*. 2004). In both cases, however, the aim is to detect the desired, low-frequency signal that reflects surface vegetation phenology while minimizing or removing the undesirable, high-frequency noise related to variable atmospheric conditions and bidirectional effects (van Dijk *et al*. 1987).

2.3.1 Function-Fitting Techniques

Of the various function-fitting strategies employed in the noise reduction of NDVI time series, the application of the Fourier transform is one of the most popular. A mathematical technique used in a variety of disciplines to separate a data set into its various frequency components (Jensen 1996) – in this case temporal frequencies – Fourier analysis is based on the concept that an infinite series of sine and cosine functions can be used to represent a function at every point in an interval (van Dijk *et al*. 1987). By generating a new NDVI time series with the appropriate temporal signal through the combination of various sine and cosine functions, Fourier-based strategies generally produce the smoothest NDVI time series.

In one of the first applications of Fourier-based techniques to noise reduction for NDVI time series, Sellers *et al*. (1994) presented a method involving a series of steps to

produce a product called FASIR-NDVI: i) Fourier wave Adjustment, ii) Solar zenith angle adjustment, iii) Interpolation of missing data, and iv) Reconstruction of NDVI data. This first step fits a Fourier series through the original NDVI time series using a robust least-squares procedure, based on the assumption that: i) NDVI times series should vary smoothly at any given point, and ii) major sources of error only decrease NDVI. Sellers *et al*. (1994) used the noise-reduced data to derive a variety of vegetative parameters for use in global biophysical monitoring efforts. Similarly, DeFries and Townshend (1994) derived global phenology-based land cover classes using NDVI time series function-fitted with Sellers *et al*.'s (1994) Fourier transform. Wang *et al* (2004) also applied this adjusted Fourier transform described in Sellers *et al*. (1994) to temporally downscale a composited NDVI data set to daily NDVI observations, whereas Wang *et al*. (2005b) employed the same method to reduce residual cloud effects and deal with missing data values in a number of multi-temporal NDVI data sets.

While the purpose of Seller *et al*.'s (1994) method is mainly for noise removal, the fast Fourier transform (FFT) technique presented in Menenti *et al*. (1993), was used for both noise removal and for helping to map differences in vegetation development (e.g. variability in growth cycle, late or early growing seasons) across Zambia. They explained that the resulting Fourier components (i.e. amplitude and phase) can be used to quantitatively describe spatial and temporal patterns of vegetation phenology (Menenti *et al*. 1993). The FFT and similar approaches have also been used in a number of other applications, including the mapping of uni-modal and bi-modal seasonal dynamics and rates of NDVI change across Africa (Olsson and Eklundh 1994), the study of vegetation phenology and classification of vegetation types across Brazil (Andres *et al*. 1994), and the mapping of vegetation in South Africa through phenological parameters (Azzali and Menenti 2000). Moody and Johnson (2001) employed a similar, discrete Fourier transform (DFT) to study inter-annual phenological variation and classify vegetation types in southern California, while Wagenseil and Samimi (2006) demonstrated the relationship between precipitation and vegetation phenology in a dry savannah environment in Namibia through the Fourier analysis of a multi-temporal NDVI data set.

Verhoef *et al*. (1996) presented an alternate Fourier-based technique called the Harmonic Analysis of NDVI Time Series (HANTS) algorithm which uses the predefined, most significant frequencies expected in a time series profile and harmonic components to apply a least-squares function-fitting procedure to the data. Unlike the above Fourier analyses, this technique does not require observations to be equidistant in time, and is therefore more flexible in its application, but does involve the setting of several parameters and more time for processing (Roerink *et al*. 2000, Verhoef *et al*. 1996). Roerink *et al*. (2000) employed the HANTS algorithm to reduce noise and error in an NDVI data set covering Europe, demonstrating the successful removal of cloud contamination. Similarly, Jakubauskas *et al*. (2001, 2002) demonstrated the application of Fourier-based harmonic analysis of NDVI time series in characterizing seasonal changes and vegetative cover change detection in the southern Great Plains of the United States.

Wavelet transforms, similar in some ways to Fourier transforms, are also recognized for their potential application to time series noise reduction (Hogan 1999). By separating a signal into several multi-resolution temporal components, they allow for the localized removal of undesirable frequencies and retention of desirable, time-varying signals (Li *et al*. 2000, Fang *et al*. 2004). Sakamoto *et al*. (2005) applied a wavelet transform technique to NDVI time series for the purpose of noise reduction and for the subsequent extraction of rice planting, heading and harvesting dates at several locations across Japan, concluding that it showed superior performance over Fourier-based time series noise reduction. Using a similar wavelet transform technique, Sakamoto *et al*. (2006) processed NDVI data to map the spatial distribution of various rice cropping systems across the Mekong Delta in south-eastern Asia. More recently, Lu *et al*. (2007) proposed an alternative wavelet-based method for noise reduction in NDVI time series which involves the use of quality flags and blue band reflectance. They found the method to perform well over a portion of eastern China when compared to alternative approaches, including a Fourier-based technique (Lu *et al*. 2007). Other wavelet transform applications to vegetation phenology include the removal of unwanted variation in investigations of El Nino/Southern Oscillation effects on inter-annual

comparisons (Li *et al*. 2000), noise reduction in ground-based spectral vegetation curves (Schmidt *et al*. 2003, 2004), the reduction of noise in NDVI time series through wavelet transforms for the detection of invasive tropical plant species across the United States (Bruce *et al*. 2006), and the wavelet-based noise reduction of NDVI time series for the prediction of seasonal foraging conditions in southern Texas (Alhamad *et al*. 2007).

In addition to Fourier- and wavelet-based noise-reduction strategies, a number of linear and non-linear regression techniques are also applied in the literature. For instance, Swets *et al*. (1999) presented a weighted least-squares windowed linear regression approach that performs a series of best-fit linear regressions for each observation in the time series, averages these and interpolates between the resulting new points to generate a continuous NDVI signal. The algorithm is upwardly biased to account for much of the noise causing decreased NDVI estimations (Swets *et al*. 1999). This method has been applied by Ricotta *et al*. (2003), who used noise-reduced NDVI time series to study the impact of climate variability on grassland species abundance in the U.S. Great Plains, and by Reed (2006), who used a noise-reduced NDVI data set covering North America to analyze 20-year trends in vegetation phenology across the continent.

In an asymmetrical Gaussian function-fitting approach to NDVI time series noise reduction, Jönsson and Eklundh (2002) fit simple local nonlinear Gaussian model functions to sets of NDVI observations. These were then combined into global functions designed to represent the full NDVI profile while preserving higher NDVI values over lower ones. Jönsson and Eklundh (2002, 2004) both demonstrated the successful extraction of phenological information from a noise-reduced time series covering the African continent, mapping the regional distribution of spring green-up dates and annual phenological amplitude. Verbesselt *et al*. (2004) used this method to help reduce noise in time series for use in the study of drought in South Africa, while Olofsson *et al*. (2007) reduce noise with the asymmetrical Gaussian function-fitting before using it to model vegetation biomass production across Scandinavia.

Beck *et al*. (2006) presented a similar function-fitting procedure, a double logistic function-fitting, whereby NDVI time series were modeled as a function of time using six

parameters: i) winter NDVI, ii) maximum NDVI during the growing season, iii) a rising curve inflection point, iv) a falling curve inflection point, v) the rate of increase at the former inflection point, and vi) the rate of decrease at the latter inflection point. This was designed for applications at high-latitude environments because it accounts for winter NDVI and the effects of snow. Beck *et al*. (2006) applied the method to an NDVI data set covering northern Scandinavia, deriving spring and autumnal phenological transition dates for the area. This method was further employed by Beck *et al*. (2007) to create an NDVI data set covering Norway, Sweden and Finland, for use in future vegetation and climate studies, including the mapping and monitoring of regional phenological transitions.

While designed more for phenological information extraction from EVI time series rather than noise reduction, Zhang *et al*. (2003) introduced a Piecewise Logistic regression function that reduces noise and smoothes time series as it extracts measures of phenology. In this procedure, periods of increasing or decreasing values are identified, and logistic functions are first fit to each period, and then combined to describe the entire EVI profile (Zhang *et al*. 2003). Zhang *et al*. (2004) employed this method to reduce noise in and detect a number of phenological transition dates across the northern hemisphere, studying their link to surface temperatures and climate variability. Pañuelas *et al*. (2004) also used this method for deriving spring green-up from time series data and linking it to patterns of rainfall and water availability in Spain. In investigating the effects of atmospheric circulation patterns on vegetation variability in Siberia, Vicente-Serrano *et al*. (2006) used the procedure described by Zhang *et al*. (2003) to process their multi-temporal NDVI data set.

Spline-based function-fitting methods have not been employed much in the literature, but were applied by Wang *et al*. (2005a, 2005b) to interpolate missing observations (removed because of poor quality) in both LAI and NDVI time series. However, the application of a new NDVI time series noise reduction technique involving a high-order spline-based function-fitting procedure was presented by Bradley *et al*. (2007). The procedure comprises a complex set of steps that include the use of harmonic and polynomial components to derive average annual phenology, and a spline fit to

describe long-term inter-annual variability in a time series, the details of which are described in Hermance (2007) and Hermance *et al*. (2007). Using this function-fitting technique, Bradley *et al*. (2007) analyzed an NDVI data set covering the Great Basin region of the U.S. and examined spring green-up dates across a number of years for various land cover types, illustrating regional inter-annual variability in phenological transition dates.

2.3.2 Filtering Techniques

As in the case of function-fitting methods, a wide variety of filtering techniques has been used to reduce noise in NDVI time series, and they range from simple moving means to more complex multi-step procedures. In an example of the former, van Dijk *et al*. (1987) presented four separate filters for noise reduction: a running weighted mean filter, a running median filter, the 4253H-twice filter and the 3RSSH-twice filter. The first two are straightforward; they simply apply a running window in which the weighted average or median of the values in the window directly before and after the value of interest, replaces the value of interest (van Dijk *et al*. 1987). These are some of the simplest filtering techniques. The third filter, 4253H-twice, applies a series of running median filters with windows of four, two, five and three, followed by a running weighted mean filter to an NDVI time series. The residuals of the filtering are then also filtered in this manner and then re-added to the filtered data, hence the 'twice' in the name of the filter (Velleman 1980). In this way, not only is noise in the time series reduced but trends in the original data are also maintained. The fourth, the 3RSSH-twice filter, iteratively runs a running median filter with a window of three to the time series until no additional changes are made, splits and interpolates resulting flat sections of the time series, then runs the weighted mean filter and same residual noise reduction step described above (Tukey 1977). While different from the 4253H-twice filter, the 3RSSH-twice filter is designed to function in a similar manner. In applying these filters to data sets covering agricultural fields in Iowa, U.S. and west central Thailand, van Dijk *et al*. (1987) demonstrated the applicability of simple filtering techniques to reducing noise in VI data sets over a variety of land cover types. Despite their simplicity and evident usefulness, these filters are not applied elsewhere in the literature for noise reduction in VI time

series, except for the 4253H-twice filter, which is used for noise reduction in other applications (e.g. Cowan and Odell 1990).

In response to the reduction in temporal resolution caused by the common application of the maximum value compositing (MVC) described by Holben (1986), Viovy *et al*. (1992) proposed a noise reduction technique for use on daily NDVI time series called the Best Index Slope Extraction (BISE). Through the use of set thresholds and a 'sliding period' the procedure searches forward within this sliding period from the first date of the time series, accepting values higher than preceding ones as long as they are within a set 'noise' threshold, and accepting a decrease only if it is followed by a period of gradual increase. In this way, it accounts not only for gradual changes in surface vegetation, but also sudden changes brought on by fire, deforestation, or crop harvest, which are followed by a period of vegetative regeneration, and can thus be differentiated from spurious drops in NDVI (Viovy *et al*. 1992). Interpolation is used to fill NDVI observations between acceptable points. Application of the BISE for reducing noise in NDVI time series also occurs elsewhere in the literature. White *et al*. (1997) used it to reduce noise in a multi-temporal NDVI data set for use in producing a continental phenology model for portions of North America, while Jia *et al*. (2004) reduced noise with the BISE in an NDVI data set used to examine phenological patterns in the Alaskan tundra, and Wang *et al*. (2004) used BISE-filtered NDVI to study the seasonally dynamic relationship between satellite- and ground-based observations. Lovell and Graetz (2001) presented a revised BISE procedure adapted for use on composited, as opposed to daily NDVI data sets. They applied this successfully to a multi-year NDVI time series covering the Australian continent.

In a more complex procedure, Chen *et al*. (2004) presented a simplified least-squares-fit convolution based on the technique described by Savitzky and Golay (1964), in which the convolution represents a weighted moving mean filter with the weights given as a polynomial of a particular degree. The method iteratively fits a new NDVI time series to the original data points until certain conditions are met (e.g. spurious lows are removed, etc.) and the resulting time series is filtered. Chen *et al*. (2004) applied this method to a multi-temporal NDVI data set covering portion of south-eastern Asia and a

variety of land cover types, reducing the effects of atmospheric conditions and bidirectional effects in the data. Jönsson and Eklundh (2004) also applied an adjusted Savitzky-Golay method to NDVI time series in Africa, which is then used to extract a variety of phenological information, including patterns of spring green-up across the continent. Though more complicated than many other filtering techniques, the Savitzky-Golay filter provides a flexible approach to noise reduction.

In a less complex attempt at removing noise, Kobayashi and Dye (2005) performed a simplified atmospheric correction on an NDVI data set using the atmospheric correction technique described in Rahman and Dedieu (1994). This technique involves the application of a set of semi-empirical formulas to describe the various interactions of solar radiation with atmospheric components, but deals only with atmospheric contamination in NDVI time series. Cihlar *et al*. (1997) proposed an alternative noise reduction procedure, referred to as the atmospheric, bidirectional and contamination corrections of Canadian Centre for Remote Sensing, (ABC3). Consisting of nine different steps, this process involves a number of atmospheric corrections, including that described by Rahman and Dedieu (1994), angular (i.e. bidirectional) corrections and the linear interpolation of missing or contaminated observations (Cihlar *et al*. 1997). Cihlar *et al*. (2004) presented a revised, improved version of the ABC3 procedure, applying the algorithm to multi-temporal data sets covering the Canadian landmass. The success of the technique was demonstrated by Cihlar *et al*. (1997, 2004), but other applications are not found in the literature. This may perhaps be due to the amount of processing and ancillary data involved in the technique.

Another filter that shows promise but does not appear elsewhere in the literature is Ma and Veroustraete's (2005) mean-value iteration (MVI) filter. Based on the assumption that each observation in the NDVI time series will closely approximate the mean of the previous and following observations, the MVI filter iteratively compares the average of the previous and following NDVI observations with the observation of interest, replacing it with this average where the appropriate conditions are met (Ma and Veroustraete 2005). In other words, where the NDVI value of interest differs too much from its neighbours, it is replaced by the mean of their values. The application of this

method to NDVI data sets covering north-western China illustrated its usefulness in minimizing time series noise (Ma and Veroustraete 2005). The basic principle behind the MVI filter is also seen in the noise reduction techniques used by Chen *et al*. (2000) and Xiao *et al*. (2002). In the former, if the NDVI at each point in the time series was lower than the mean of the observations immediately before and after it, it was replaced with this mean; in the latter, dates identified as contaminated by cloud were replaced by the mean of the preceding and succeeding NDVI values. While Chen *et al*.'s (2000) method may deal with both atmospheric and bidirectional effects, only atmospheric effects are accounted for in Xiao *et al*.'s (2002) method.

Filipova-Racheva and Hall-Beyer (unpublished) presented a number of autoregressive moving window algorithms as a means to reduce noise in NDVI time series, most consisting of a combination of autoregressive moving average (ARMA), autoregressive moving maximum (ARMM) and/or an autoregressive moving median (ARMD) algorithms. The combinations included: i) an ARMA algorithm with a) a three-date window or b) a five-date window; ii) an ARMM algorithm in combined with an ARMA with a) a three-date window, b) a five-date window, or c) a five-date window applied twice; and iii) an ARMD with a three-date window, followed by an ARMA with a five-date window. Although the work is unpublished, Filipova-Racheva and Hall-Beyer (unpublished) demonstrated the appeal of these simple types of algorithms, as do van Dijk *et al*. (1987), with the 4253H-twice and 3RSSH-twice filters. Not only are they easy to implement, but they can be effective at removing the unwanted noise in NDVI time series.

An alternative filtering technique was presented by Kang *et al*. (2005), who applied a temporal and spatial interpolation scheme to cloud-contaminated pixels in a multi-temporal data set. While the application of this technique did not actually involve NDVI time series, it could be applicable to such data sets. The method comprises a spatial interpolation first, in which pixels identified as cloudy for a particular date are replaced with a new interpolated value from surrounding cloud-free pixels (Kang *et al*. 2005). If no cloud-free pixels are available for that date, for that pixel, temporal interpolation from previous cloud-free or interpolated dates is performed. Kang *et al*.

(2005) successfully applied this method to data sets covering a portion of the north-western U. S., but concede that it is only applicable in cases where sufficient cloud-free data exist for the interpolation.

Knight *et al*. (2006) described a multi-step noise reduction process for MODIS NDVI time series that involves both filtering and function-fitting techniques. First, anomalous spikes or drops in the time series are identified using an empirically-derived threshold value, designed to detect sudden increases or decreases that are followed by a return to values near those of previous data points; these are removed and the data points flagged. Second, using MODIS quality flags, data values of unacceptable quality are identified, removed, and flagged (Knight *et al*. 2006). In the third and final step, a Fourier transform is used to separate noise from the desired temporal signal in the data and produce a fitted curve, the values of which are used to replace the deleted and flagged values from previous steps. The resulting time series are then applied in a multi-temporal land cover classification of a study area covering eastern portions of North Carolina and Virginia. By incorporating both filtering and function-fitting techniques in their process, Knight *et al*. (2006) were able to exploit the advantages of both types of methods, while minimizing their respective disadvantages. However, the setting of initial thresholds and a minimum quality flag for selecting pixels for replacement can be highly subjective.

2.4 Extracting Phenological Information: Time Series Metrics

Some of the first regional- to global-scale applications of satellite remote sensing to studies of vegetation phenology involved qualitative assessments of the spatial and temporal patterns of seasonal vegetative growth and development. For example, Justice *et al*. (1985) described global phenology in broad terms through a comparison of four NDVI composite images covering portions of April through May, June through July, September through October, and December through January, observing transitions in vegetative developmental phases over various regions. Townshend and Justice (1986) visually compared monthly NDVI composites over Africa between 1984 and 1985, demonstrating considerable continental variation in phenology across seasons and between years.

Such qualitative assessments were soon replaced, however, by more quantitative methods, most of which involve the derivation of metrics from per-pixel time series of NDVI, designed to identify and measure phenological transitions. Traditional ground-based measures of phenology involve the observation of individual plants or groups of individual plants of one species and identify specific events such as budburst, leaf-out, or flowering (Badeck *et al*. 2004). However, the nature of satellite imagery (i.e. its spatial and temporal resolutions) results in a measure of aggregate vegetative growth and development over a much larger area, so that although phenological dynamics are captured by the sensor, these events do not reflect characteristics of plant species, populations or communities but rather, the ecosystem level characteristics of an individual pixel (White *et al*. 2002, Reed *et al*. 2003). Because it is assumed that gradual, yearly variation in NDVI is directly related to levels of surficial vegetative photosynthetic activity, it is therefore also assumed that broad measures of phenology and the related changes in photosynthesis can be derived directly from NDVI temporal profiles. As mentioned above, this information is often extracted in the form of metrics, derived from NDVI time series. Such metrics take a variety of forms and can bear a constant relationship to the timing of particular events or seasonal transitions and the associated levels of photosynthesis, as well as overall seasonal characteristics of the yearly growing season and vegetation development.

2.4.1 The Start of Season Metric

One of the most prevalent NDVI time series-derived metrics concerns the study of spring phenology, namely, the timing of the start of the growing season, also referred to as start of (spring) season (SOS), onset of greenness, onset of green-up, green-up, and the green wave (Schwartz 1998, Schwartz *et al*. 2002, Kang *et al*. 2003, Zhang *et al*. 2003). It will be referred to here as SOS (Table 2-3). The determination of SOS is in important step in estimates of ecosystem growing season productivity (e.g. White *et al*. 1999), is sensitive to climate variability and can be an important component of climate research (e.g. Schwartz and Reiter 2000, Menzel *et al*. 2001, Ahas *et al*. 2002, Ho *et al*. 2006), and is closely linked to terrestrial-atmospheric ecosystem mass and energy exchanges, particularly CO_2 fluxes (e.g. Hall *et al*. 1975, Chen *et al*. 1999, Wilson and Baldocchi

2000). It has been of considerable interest in ecological, phenological, and particularly, climatological research.

A number of studies applying NDVI-derived SOS for studying long term trends in phenology have shown an advance in the timing of this event across northern latitudes. Myneni *et al*. (1997, 1998) estimated a general advance of 8 ± 3 days in SOS between 1981 and 1991 in regions north of 45^0 latitude using global NDVI data, observing the most change in regions of high vegetative biomass. Zhou *et al*. (2001), in a similar analysis, reported advances in SOS of 8 ± 3 and 6 ± 2 days between 1982 and 1999 at latitudes of 40^0N to 70^0N across North America and Eurasia, respectively, contrasting the fragmented patterns of change across North America with the more homogeneous patterns seen over large swaths of land across Eurasia. Tucker *et al*. (2001) also described advances in SOS at latitudes of 45^0N to 75^0N of 2 ± 1 days on average between 1992 and 1999, while Delbart *et al*. (2006) reported an advance of 8 and 3.6 days between 1982 and 1991, and 1993 and 2004, respectively, for northern Eurasia. In many of these studies, the linkage between earlier timing in the start of spring plant growth and increasing average spring temperatures was suggested.

In addition to the study of temporal trends in the timing of SOS, this satellite-derived phenological measure has also been used in conjunction with additional variables in land cover classification (e.g. DeFries *et al*. 1995, Wagenseil and Samimi 2006), testing meteorologically-based phenological models (White *et al*. 1997), studying urban heat island effects on urban phenology (White *et al*. 2002), investigating the relationship between grass carbon pathway distribution and climatic factors (Ricotta *et al*. 2003), and in differentiating tundra types in Alaska using phenology (Jia *et al*. 2004).

Despite the importance of SOS as a phenological measure, however, no standard technique for deriving the SOS metric exists, and that a variety of methods is used in its generation from multi-temporal NDVI data sets. Reed *et al*. (2003) divided this range of methods into three broad types: i) threshold techniques, ii) inflection point techniques and iii) curve derivative techniques. The first involves both absolute and relative NDVI thresholds at which, once reached by the spring values in an NDVI profile, the growing season is determined to have begun. Based on the idea that SOS indicates the point at

which spring plant growth has become sufficiently dense for detection by the satellite sensor over background conditions, these methods aim to capture this detection (Jia *et al*. 2004). For instance, Justice *et al*. (1986) presented one of the earliest applications of the threshold technique, where the presence of photosynthetically active vegetation at the SOS in semi-arid Africa was identified using an NDVI threshold of 0.09. Lloyd (1990), Fischer (1994a), Suzuki *et al*. (2003), and Jia *et al*. (2004) also applied thresholds (i.e. 0.099, 0.170, 0.200 and 0.090, respectively) in deriving SOS over global to regional scales, for the purposes of land cover and crop classifications. The use of absolute thresholds such as these can be problematic, however. As Reed *et al*. (1994) explained, thresholds vary with vegetation type, illumination conditions and soil background, and are therefore not applicable over all global regions or land cover types, nor can they generally be extended beyond the region over which they were derived. Relative thresholds are a more flexible alternative for deriving SOS.

Unlike absolute thresholds, relative thresholds are more widely applicable, are less dependent on local or regional conditions, and are more commonly found in the literature. White *et al*. (1997), for example, used an annually-defined $NDVI_{ratio}$ threshold of 0.5, where $NDVI_{ratio}$ refers to the proportion of total greenness reached at a particular date in the NDVI profile, and is calculated as the difference between the NDVI value of interest and the annual minimum NDVI, divided by the difference between the annual maximum and minimum NDVI values. A threshold of 0.5 reflects the point at which 50% of annual total greenness is reached. White *et al*. (2002) used a modified version of this method, a seasonal-midpoint NDVI (SMN) relative threshold, for identifying SOS, which sets the threshold as half the yearly NDVI amplitude (i.e. the difference between annual maximum and minimum NDVI). The SMN threshold method was also employed by Badeck *et al*. (2004) to NDVI data covering homogeneous deciduous forest stands in Europe, by Bradley *et al*. (2007), who employed SOS in the differentiation between natural land cover variability and land cover change, and by Fisher *et al*. (2006), who attempted to scale between ground- and satellite-based phenological observations. Jönsson and Eklundh (2002) again used a similar method for defining SOS, as 10% of the distance between pre-spring minimum NDVI and maximum NDVI. This was also

applied by Jönsson and Eklundh (2004), who used it to map SOS across the African continent. Delbart *et al*. (2006) used a different type of relative thresholding approach, whereby a pixel-specific NDVI threshold was set based on the average NDVI value observed at the time when the normalized difference water index (NDWI) had begun to increase after snowmelt. This method was designed to identify SOS in high latitudes as the point at which snow has melted and vegetation had begun to grow, and is indicated by an increase in NDWI after a preceding decrease, assumed to reflect snowmelt in the spring.

While absolute and relative thresholds are well-used in the detection of SOS, they are set using a wide variety of criteria, as evidenced by the examples given above. The second category of techniques to derive SOS, inflection points (Reed *et al*. 1994), involves more consistent methods for generating this metric. Inflection points are those points at which the curvature of a fitted-function changes (Zhang *et al*. 2003), and in the case of NDVI time series, the point at which the profile shifts from a neutral to a positive slope in spring, which is assumed to reflect the beginning of the growing season, or SOS. Although the technique is not applied to NDVI time series by Badhwar (1984), the author presented one of the earliest uses of inflection points in deriving SOS from a polynomial function fit through a temporal profile of vegetative greenness values, which was then used in discriminating crop types in the northern Great Plains of the U.S. Moulin *et al*. (1997) also derived SOS as the point between the first and second derivatives of a temporal profile, but applied the technique to NDVI time series on a global scale and the characterization of general land cover types. In an analysis of inter-annual variation of phenology in grass- and shrublands of North America, Paruelo *et al*. (1998) defined SOS as the first inflection point in a double logistic function fit through NDVI time series, following the method outlined in Fischer (1994b). They were able to link climatological characteristics of temperature and precipitation to this and other phenological metrics over the study area. Similarly, Zhang *et al*. (2003, 2004) applied a Piecewise Logistic function of time to multi-temporal NDVI data sets while deriving SOS as the first point of maximum curvature in the function. While mapping the timing of SOS across New England and the northern hemisphere, respectively, Zhang *et* al. (2003, 2004)

demonstrated the use of the method for the large-area monitoring of vegetation using MODIS NDVI products. This method was also employed by Ahl *et al*. (2006), who derived SOS for the comparison of various satellite-based vegetation index products including NDVI, with ground-based observations, and demonstrated a consistently earlier detection of SOS by the satellite sensor. In a final example, Beck *et al*. (2006) derived SOS using inflection points from both Double Logistic and Asymmetric Gaussian functions, in a comparison of these methods for the processing of multi-temporal NDVI data. They found that both performed fairly similarly in predicting SOS, though some differences were also detected.

While inflection points reflect times of transition between curve derivatives in an NDVI time series, the third type of method for deriving the SOS metric from these time series relies on the characteristics of curve derivatives themselves. These methods are designed to identify a rapid, sustained increase in an NDVI profile, detecting the greatest rate of vegetative growth in spring as an indication of SOS (Reed *et al*. 2003). The delayed moving average (DMA) method described by Reed *et al*. (1994) for deriving metrics, including SOS, from NDVI time series is one such technique. By comparing a time-lagged, noise-reduced version of the time series produced by a DMA to the original time series, the authors identified departures from a trend in the profile. The point at which the DMA profile crosses the original profile in an upward direction and remains above it, reflects the period of greatest sustained increase, and is therefore identified as SOS (Reed *et al*. 1994). This method was also applied by Schwartz and Reed (1999), who compared the derivation of SOS dates from satellite NDVI data with climatologically-based ground models of phenology and SOS, and by Wang *et al*. (2004), in a study of the relationship between satellite and ground-based estimations of NDVI. In an analysis of phenological trends across North America between 1982 and 2003, Reed (2006) identified regions of advance, and in some cases, delay in the timing of SOS, using the DMA technique, while Lee *et al*. (2002) applied this technique for deriving SOS in a region covering inner Mongolia, and reported differing results in multi-year SOS timing trends for different land cover types. Yu *et al*. (2003) employed a modified version of Reed *et al*.'s (1994) DMA method, incorporating constraints on the selection

of SOS to within a certain portion of the year and above a particular NDVI threshold in a study of SOS across inner Mongolia, while Ricotta *et al*. (2003) employed a method very similar to the DMA technique for deriving SOS that uses a moving average window to create a time-lagged temporal profile used to identify SOS, in their investigation of grass species abundance across the Great Plains.

In contrast to the more complex methods of inflection point-based SOS derivation used by Reed *et al*. (1994) and others, DeFries *et al*. (1995) simply used the time of greatest increase between consecutive NDVI composite observations to define SOS as one of several metrics used for improving global land cover classification by the incorporation of dynamic phenological input. Piao *et al*. (2006) applied a different curve derivative technique to identify SOS by detecting the greatest rate of change between the NDVI of each date and that of the date before it through a relative index. The authors generated what they refer to as $NDVI_{ratio}$ by dividing the difference between the NDVI value of interest and the one before it (in the time series) by the NDVI value of interest. This differs from the $NDVI_{ratio}$ derived by White *et al*. (1997). The $NDVI_{ratio}$ derived by Piao *et al*. (2006) more closely reflects a rate of change in NDVI between successive composite periods, and the maximum $NDVI_{ratio}$ was used by Piao *et al*. (2006) to detect SOS in a study of phenological changes across China. The authors were able to identify temporal trends in the timing of SOS using this method.

With such a wide variety of methods for deriving SOS occurrence and timing in the literature, it is not surprising that a number of studies have shown that the general phenological event reflected by remotely-sensed NDVI-derived SOS varies with the method used to detect it. For example, in a comparison of Reed *et al*.'s (1994) DMA method and White *et al*.'s (1997) SMN method for deriving SOS, Schwartz *et al*. (2002) found that the DMA technique produced consistently earlier SOS estimates than the SMN technique. While DMA-derived SOS detects the first sustained increase in springtime NDVI, therefore reflecting the first flush of surface vegetative greenness in understory species, SMN-derived SOS identifies the point at which half of the maximum surface greenness is reached, and was shown to reflect initial leaf expansion in dominant overstory vegetative species (Schwartz *et al*. 2002). In this way, the two methods of

deriving the SOS metric are measuring fundamentally different phenological events or processes. Similarly, Reed *et al*. (2003) argued, using a theoretical comparison of the DMA, SMN, an inflection point method and a curve derivative method, that inflection point methods would generally provide the earliest estimates of SOS, followed by curve derivative techniques, and finally, by thresholding techniques. This might differ if an alternative set of techniques were compared, but these alternative comparisons would doubtless still show distinct differences in the resulting SOS estimates.

Despite the evident importance of understanding the relationship between the methods used for deriving SOS dates from NDVI time series, and the corresponding phenological processes occurring on the surface, few comparisons like those described above have been made. Nevertheless, it is clear that an understanding and recognition of the implications is required when selecting a method for deriving the SOS metric in the remote sensing of vegetation phenology. It is also important to note that these concerns apply not only to SOS, but to the multitude of additional metrics that are derived from NDVI time series. Though not as popular as SOS these allow for the extraction of useful, complementary phenological information, and are further discussed below.

2.4.2 End of Season, Length of Season and Other Metrics

While the SOS metric is undoubtedly the most popular remote sensing-based measures of phenology, it is commonly associated with two other, closely related metrics: the end of the growing season (EOS), also known as end of season, end of greenness, greenness offset and onset of dormancy (Reed *et al*. 1994, White *et al*. 1997, Zhang *et al*. 2003), and the length of the growing season (LGS) (Table 2-3). EOS is generally derived in a manner similar to SOS, and is itself, not often of great interest, but is more frequently used for the estimation of LGS, simply defined as the length of time between SOS and EOS (Reed *et al*. 2003). However, several studies have investigated temporal trends in phenology through changes in EOS with time, though these are often in combination with SOS and LGS. For example, Myneni *et al*. (1997) reported a delay in EOS by ± 2 days between 1983-83 and 1989-90, contributing to a lengthening of the growing season north of 45^0 latitude of 12 ± 4 days over the 1980s, while Zhou *et al*. (2001) showed delays in EOS of 4 ± 3 and 11 ± 3 days between 1982 and 1999 in North America and Eurasia,

respectively, contributing to increases in LGS by 12 ± 5 and 18 ± 4 days over each of the two continents. Much of the research involving SOS, EOS and LGS is focused on linking temporal trends in the timing of these events and climatic variation and change (e.g. Myneni *et al.* 1997, White *et al*. 1997, Ricotta *et al*. 2003, Badeck *et al*. 2004, Reed 2006). While trends in a delay of EOS have been linked with an increase in LGS, it is often suggested that much of the lengthening in the growing season is due more to advances in SOS than delays in EOS, as SOS in many regions is more closely dependent on temperatures and EOS, on photoperiod (e.g. White *et al*. 1997, 2002).

The SOS, EOS and LGS metrics are prevalent in the literature concerning the remote sensing of vegetation phenology. Not only are they important for the parameterization of surface phenology through remotely-sensed NDVI time series as evidenced by the above discussion, but they also form a foundation for the calculation of additional metrics. For instance, the integrated NDVI (I-NDVI), often calculated as the area under the NDVI curve during the growing season (Table 2-3), is of considerable interest to researchers and has been linked with annual NPP and seasonal biomass production (Tucker *et al*. 1985, Goward *et al*. 1985, Prince 1991). Justice *et al*. (1985) compared spatial patterns of satellite-derived I-NDVI over eastern Africa with ecoclimatic zone maps of the area, finding general correspondence between the two. In their comparison of grass- and shrubland vegetated sites with inter-annual climatic variation, Paruelo *et al*. (1998) used I-NDVI to show lower inter-annual variability in biomass production during the growing season than in the timing of the growing season. Reed (2006) demonstrated trends of increasing I-NDVI over particular regions of North America, concluding that changing land-use practices have significantly influenced vegetation phenology and seasonal primary production in the Canadian Prairie Provinces.

In addition to the SOS, EOS, LGS and I-NDVI metrics, including the level of NDVI and timing at which the first two occur, a variety of other less common, though informative metrics are also found in the literature (Table 2-3). One such metric, the timing and value of the maximum seasonal NDVI, reflects the maximum level of photosynthetic activity reached during the growing season, and was used by Reed *et al.* (1994), DeFries *et al.* (1995) and Paruelo *et al.* (1998) as one of several metrics derived

from NDVI time series for application to land cover classification and in the study of inter-annual variation in vegetation phenology. This metric and the minimum seasonal NDVI were used by White *et al*. (2002), Jönsson and Eklundh (2002, 2004) and Bradley *et al*. (2007) to derive NDVI amplitude (Table 2-3) – the difference between minimum and maximum NDVI during the growing season – as well as rates of spring green-up and autumnal senescence (Table 2-3), commonly calculated as the general slope between SOS and maximum NDVI, and between maximum NDVI and EOS, respectively (Reed *et al*. 1994, DeFries *et al*. 1995). The former corresponds to the seasonal amplitude of vegetation greenness, and can vary considerably between vegetation types, while the latter reflects the asymmetry of a growing season NDVI profile (e.g. rapid green up vs. slower period of senescence). Additional metrics are listed in Table 2-3.

Table 2-3: NDVI time series metrics, derived for the characterization of surface vegetation phenology.

Metric[1]	Description	Methods of Estimation
Start of Growing Season (SOS)	Timing and level of photosynthesis at the start of measurable photosynthesis, representing the start of the growing season	Thresholds, inflection points, curve derivatives, or Fourier-based methods
End of Growing Season (EOS)	Timing and level of photosynthesis at the cessation of measurable photosynthesis, representing the end of the growing season	Thresholds, inflection points, curve derivatives, or Fourier-based methods
Length of Growing Season (LGS)	The duration of measurable photosynthesis, representing the length of the growing season	Number of composite periods or days between SOS and EOS
Mid-Season	Timing and level of photosynthesis at the middle of the growing season	Point midway between time at which 90% of the seasonal amplitude is reached and time at which 10% of senescence has occurred
Integreted NDVI (I-NDVI)	Overall productivity and biomass produced during the growing season	Area under the NDVI curve between SOS and EOS
Rate of Green-Up	Speed at which spring green-up occurs	Slope of line between SOS and maximum NDVI
Green-Up Fraction	Portion of growing season spent in green-up	Number of days/composite periods between SOS and maximum NDVI over the total LGS
Green-Up I-NDVI	Portion of total growing season productivity produced during green-up	Area under the NDVI curve between SOS and maximum NDVI, over I-NDVI
Start of Senescence	Timing and level of photosynthesis at the start of rapid decrease in photosynthesis	Inflection point method
Rate of Senescence	Speed at which autumnal senescence occurs	Slope of line between maximum NDVI and EOS
Senescence I-NDVI	Portion of total growing season productivity produced during senescence	Area under the NDVI curve between maximum NDVI and EOS, over I-NDVI
Maximum NDVI	Timing and level of photosynthesis at maximum photosynthesis reached during the growing season	Maximum NDVI between SOS and EOS
Minimum NDVI	Timing and level of photosynthesis at minimum photosynthesis reached during the growing season	Minimum NDVI between SOS and EOS
Mean NDVI	Overall mean level of photosynthesis occurring throughout the year	Mean of NDVI values over a year
NDVI Amplitude	Range in levels of photosynthesis occurring during the growing season	Range between maximum and minimum NDVI
Modality	Number of growing seasons within a year	Periodicity of significant peaks in NDVI time series
Relative NDVI Range	A standardized NDVI amplitude, used in inter-annual comparions of biomass productivity	NDVI Amplitude divided by I-NDVI
Harmonic Amplitude	Similar to NDVI Amplitude	Wave amplitude of harmonics derived from Fourier transforms
Harmonic Phase	Timing of maximum photosynthesis	Wave phase of harmonics derived from Fourier transforms

[1]For more details on these, see Reed et al. (1994), DeFries et al. (1995), Azzali and Menenti (2000), Jönsson and Eklundh (2002, 2004), and Pettorelli et al. (2005).

While the above metrics are either based primarily on the derivation of SOS and EOS for their calculation, or on thresholds, inflection points or curve derivatives, further sets of metrics have been produced from multi-temporal NDVI data sets using Fourier transform-based techniques. Such techniques are based on the assumption that the sum of a series of cosine waves (also referred to as 'terms') with an additive term, can be used to parameterize the characteristics of a complex curve (Jakubauskas *et al*. 2001). As explained in the previous section, these are used in the reduction of noise for NDVI time series, but they can also provide a means of describing and identifying surface phenological processes from such time series. In much of the literature amplitude and phase angle – the height and offset from point of origin for a wave, respectively (Table 2-3) – are used in the characterization of seasonal phenology (i.e. they reflect the amplitude of photosynthetic activity during the growing season, and the timing of this activity, respectively). For example, Menenti *et al*. (1993) applied a fast Fourier transform (FFT) to time series of NDVI, extracting amplitude and phase in order to map the timing and level of vegetative growth across Zambia, and to map zones of similar vegetative phenological patterns. Azzali and Menenti (2000) used this same technique and the resulting amplitude and phase information, to classify vegetation-soil-climate units across southern Africa, illustrating the relationship between aridity, vegetation type and plant phenology in this region.

While using a slightly different set of phenological measures (e.g. maximum green-up, maximum and minimum NDVI, etc.) in their application of a Fourier-based technique to NDVI time series, Olsson and Eklundh (1994) nonetheless demonstrated the mapping of phenological uni- and bi-modal phenological cycles across the African continent. Verhoef *et al*. (1996) mapped phase and amplitude values across South America derived from HANTS, using various patterns of phase and amplitude to describe spatial patterns in phenology over the region. In their application of a DFT technique to a small portion of southern California, Moody and Johnson (2001) linked phenological profiles with basic vegetation formations, producing an unsupervised classification of these formations over the study area. Amplitude and phase derived from Fourier-based analysis of NDVI time series was also used by Jakubauskas *et al*. (2001), who employed

these metrics for phenology-based crop-type identification in south-western Kansas. In an effort to link spatial variation in vegetation phenology with inter-annual rainfall variability in a park in northern Namibia, Wagenseil and Samimi (2006) also employed amplitude, phase and mean NDVI derived using similar Fourier-based techniques. Indeed, while these Fourier-derived metrics are not often combined with the SOS, EOS and other metrics described above in their application to various research objectives, their applicability to the study of vegetation phenology through remotely-sensed NDVI time series is evident.

2.5 Chapter Summary

The remote sensing of vegetation phenology is a diverse field of research, comprising a multitude of disciplines, research objectives and approaches to studying the health and development of surface vegetation as reflected in the occurrence and timing of periodic biophysical events. The above chapter reviews various aspects of this topic, presenting an overview of the relevant literature concerning the use of satellite imagery in the multi-temporal study of surface vegetation. The importance of vegetation phenology to a variety of biological, ecological and climatological concerns is evident in the literature, as is the need for repeatable regional- to global-level data sets in support of large-area studies of earth system dynamics and processes, as is provided by a variety of satellite-borne optical sensors. MODIS in particular, is becoming increasingly popular for phenological research. While it provides data at spatial and temporal resolutions similar to that of the AVHRR, a popular long-standing sensor, it also provides improved radiometric correction, as well as a multitude of additional vegetation related products. The MODIS VI data sets are especially interesting to those who study vegetation phenology as these are shown to be closely related to levels of photosynthetic activity occurring at the surface. Of the two VIs, the NDVI is particularly prevalent throughout the literature. The simplicity of its calculation, its normalizing nature, and its availability as pre-processed data sets, such as those provided by MODIS, have rendered it very popular for phenological research.

A MODIS NDVI vegetation product provides the remotely-sensed data set for the current research. The free and easy access to MODIS data sets through the LP DAAC

and the superior radiometric quality of these data to the more common AVHRR data sets led to the selection of MODIS VI data. In addition, the continued popularity of the NDVI rendered it the most practical choice for providing the most substantial contribution to present-day research and literature regarding the remote sensing of vegetation phenology. Nevertheless, the inherent noise and error resulting from high-frequency variations in atmospheric conditions and sun-sensor-surface viewing geometries that is present in the time series of NDVI so widely applied in phenological research are an important concern.

It is the presence of noise in multi-temporal NDVI data sets, and the importance its minimization to subsequent efficient use in studying vegetation phenology, which provides the impetus for the current research. Numerous researchers have addressed this problem, providing a variety of noise reduction filtering and function-fitting techniques that work on a per-pixel basis for the minimization of undesirable high-frequency noise. However, no standard methods or approaches prevail in the literature and few comprehensive studies testing the relative success of many of these techniques are found. While it is obvious from the present review that a number of such techniques demonstrate considerable potential for noise reduction in NDVI time series, rarely are these techniques applied elsewhere beyond their initial proposals. The present project aims to provide a more comprehensive comparison of several of the noise reduction techniques presented above, and thus relies heavily on this portion of the present literature review. Additional information is provided regarding more peripheral aspects of the remote sensing of vegetation phenology, particularly in relation to the current research, but these portions of the literature review are relevant to a comprehensive understanding of the topic and of the methods used to undertake the present empirical comparison.

The final element of the present review concerns the extraction of phenological information in the form of metrics from multi-temporal remotely-sensed NDVI data sets. Although such metrics do not form the main component of the present research, knowledge of their importance and application in the remote sensing of vegetation phenology is pertinent to this research. They comprise an important aspect of the empirical comparison described in the following chapter, and therefore require consideration in the above literature review. In particular, the extraction of start, end and

length of growing season metrics are important phenological measures, playing integral roles in Earth-atmosphere energy and mass exchanges and climatological studies because of their responsiveness to local and regional environmental conditions. Deriving these metrics from multi-temporal NDVI data sets involves per-pixel time series calculations, but is done using a variety of methods including thresholding and inflection points. Such metrics continue to support numerous biological, ecological and climatological research objectives and provide a method for studying, measuring and monitoring short- to long-term variations in vegetation phenology at regional to global scales.

The present literature review provides a summary of the current knowledge concerning the remote sensing of vegetation phenology. It offers a summary of those topics pertinent to the current research and of related topics which, though perhaps not directly essential, are certainly relevant to an understanding of the procedures and methods described in the following chapter.

Chapter Three: Methods

The following chapter describes the methods used to complete an empirical comparison of selected strategies for the noise reduction of NDVI time series. Figure 3-1 presents a flowchart outlining these methods and the principal steps involved. The analysis was undertaken within a model framework. More specifically, the application of selected noise reduction techniques was tested on a series of model NDVI time series with introduced noise, rather than actual NDVI time series. The decision to employ a model environment was based on a consideration of the difficulties involved in acquiring and applying the ground truth data that would have been necessary for a real-world analysis of the selected noise reduction techniques. For instance, the current investigation would require not only the collection of surface estimates of NDVI, but also vegetation phenology. Collection of either set of data was not feasible within the scope of the current project. Not only would the collection of simultaneous red and near-infrared reflectance data over an area of 250 meters (m) by 250 m – the size of one MODIS pixel – have been necessary, but so would the collection of ground observations of surface phenology over the same area for many months. The collection of such data is very costly, requiring not only the appropriate equipment but extensive field work, and would also ideally be done over a number of such plots in order to provide ground truth for more than one MODIS pixel.

In addition to difficulties in acquiring sufficient ground-based NDVI and phenological data, scaling up these data to provide appropriate ground truth for MODIS pixels would have been problematic. Current validation efforts by MODIS Land team require a number of air- and satellite-borne data sets of varying spatial resolutions in order to link *in situ* ground observations with MODIS pixel reflectance (Morisette *et al*. 2002, Gao *et al*. 2003). Such efforts were beyond the range of the current research, and in terms of ground-based phenological observations, data concerning one or even several vegetative species over a 250 m by 250 m area may not have corresponded to what is observed in a remotely-sensed NDVI time series profile – these profiles represent an agglomeration of all the phenological processes occurring at the surface, and include not only canopy vegetative health and development, but that of the undergrowth as well. A

MODIS pixel does not capture species, population or even community-level phenological characteristics, but rather, those of all surface elements found within one pixel (White *et al*. 2002, Reed *et al*. 2003). In light of these difficulties, a model environment was chosen as a viable, practical, and cost-effective approach to the present investigation of noise reduction in NDVI time series.

A set of model NDVI time series reflecting ideal conditions and constructed from a MODIS NDVI data set covering the study area for 2003 through 2005, formed the basis of the current investigation (Figure 3-1). Several levels of noise were introduced to these model time series, affording a variety of modeled noisy NDVI time series on which the six candidate noise reduction techniques, selected from the literature, were tested. It was the goal of the present research to test the quality of these noise reduction techniques in terms of both i) their ability to return noisy NDVI measurements to their 'true' values and ii) their ability to maintain the integrity of the original NDVI signal for the subsequent extraction of phenological metrics. The empirical comparison of the six techniques therefore comprised two separate evaluations: one involving root mean square error (RMSE) and the other, phenology-based metrics derived from NDVI time series. Each of these evaluations provides a different measure of noise reduction technique performance, and each is equally important to a comprehensive understanding of the abilities of the selected techniques to minimize noise in NDVI time series and to maintain the integrity of the original signal. RMSE provides a means of investigating the general ability of a noise reduction technique to minimize noise in an NDVI time series by measuring the average deviation between two sets of values. That is, it allows for the assessment of how well a noise reduction technique is able to return a noisy time series to it original modeled state. However, because RMSE is simply an average measure, it only very indirectly accounts for how well the shape and amplitude of the original NDVI signal is preserved when noise reduction is applied.

NDVI time series metrics, on the other hand, are much more responsive to the shape and amplitude of a time series signal than general RMSE calculations. An investigation of the derivation of these metrics from noise-reduced NDVI time series provides a clearer measure of how well the integrity of the original signal is maintained

by the candidate noise reduction techniques. A shift in the amplitude or shape of the original signal will result in differing time series metric calculations that reflect this shift.

Combining RMSE and metric-based assessments of noise reduction technique performance provided a more comprehensive evaluation of these techniques and their application to NDVI time series. By investigating not only the general noise-minimizing abilities of the selected techniques but also, their respective abilities to preserve the integrity of the original NDVI signal, a clearer understanding of their application and use in the remote sensing of vegetation phenology could be more fully achieved.

As indicated in Figure 3-1, the results of both the RMSE and metric evaluations were transformed into performance scores, which allowed for the direct comparison between different types and categories of results that would not have been possible otherwise. These performance scores were further used in scenario summations, which comprised an assessment of how often and in what circumstances the application of a noise reduction technique to noisy NDVI time series produced a better result than when such a technique was not applied. Overall and stratified results were produced from this analysis, the latter representing the re-organization of these results into land cover, noise level and other categories in order to evaluate the effects of this stratification on the results and their reflection on the performance of the selected techniques (Figure 3-1). In addition to this, further analysis of RMSE and metric performance scores involved the generation of standardized performance scores, from which summary performance scores could be calculated. The latter allowed for the subsequent overall comparison and ranking of the six candidate noise reduction techniques and their respective performances, while stratification of these results provided for additional investigation into the effects of land cover, noise level and other factors on technique performance.

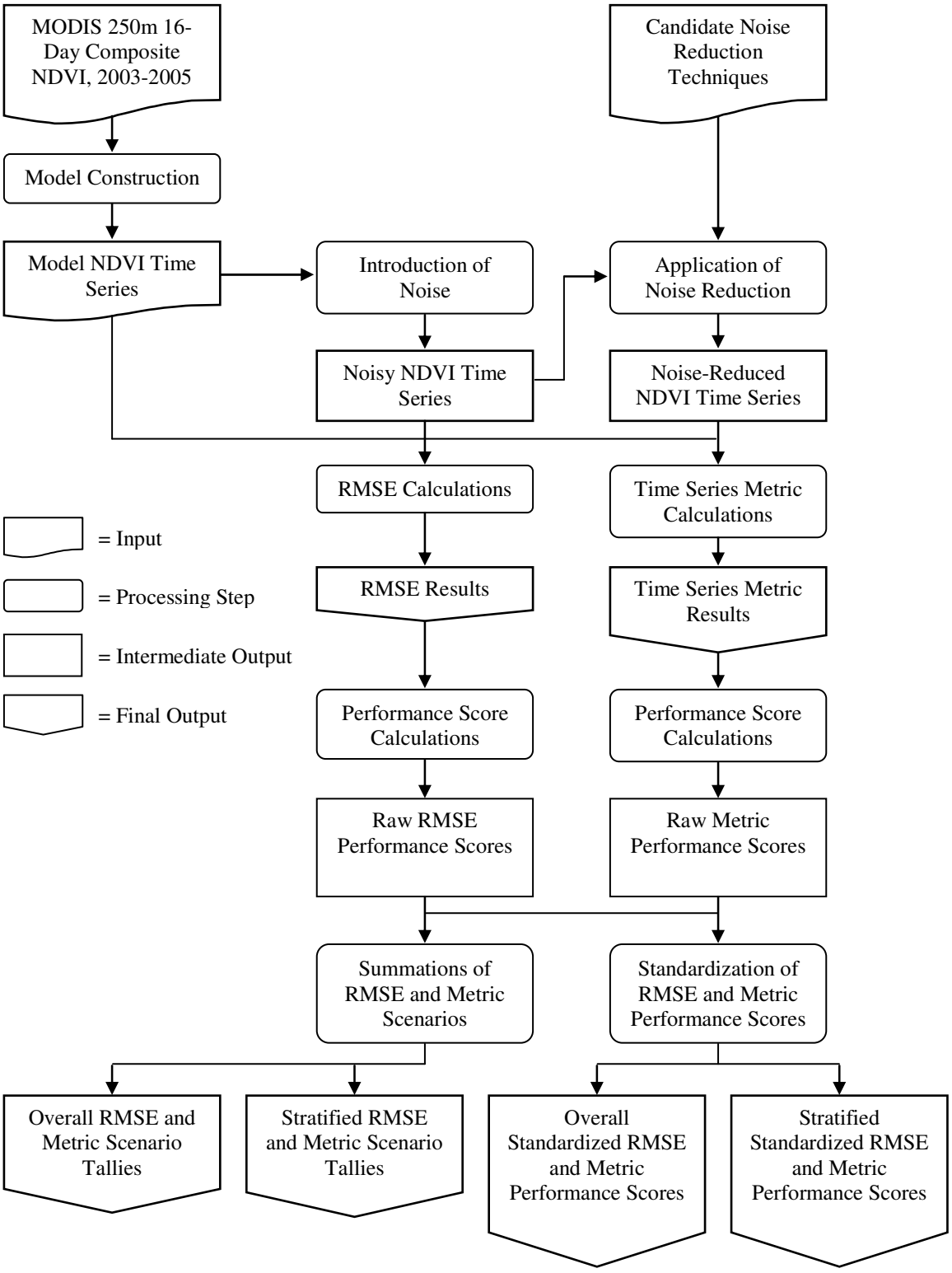

Figure 3-1: Flowchart of the principal steps involved in the empirical comparison of the selected candidate noise reduction techniques.

The procedures and steps summarized in Figure 3-1 are each described in the sections and subsections presented below. First, a description of the study area, located in west-central Alberta, Canada, is given in section 3.1, followed by explanations of the acquisition and pre-processing of the multi-temporal NDVI data set employed in the current research. Section 3.3 describes the construction of the model environment, including the generation of model NDVI time series and the introduction of noise to these time series. The selection of the six candidate noise reduction techniques and their respective implementations are presented in section 3.4, while sections 3.5 and 3.6 explain the completion of the two components of the empirical comparison, the RMSE and metric-based evaluations. Analysis of the data is described in section 3.7, including the calculation of raw and standardized performance scores, which provided a means of assessing the performance of the six noise reduction techniques, particularly in relation to one another, and to the noisy, uncorrected data.

3.1 Study Area

The current research concerns a portion of west-central Alberta, Canada, along the front ranges of the Rocky Mountains, and covers approximately 71,500 square kilometres (Figure 3-2). The reason for this choice of study area is principally due to present and future collaborations with members of the Foothills Model Forest Grizzly Bear Research Program (FMFBGRP). The study area illustrated in Figure 3-2 represents the portion of the current FMFGBRP study area that is covered by one MODIS tile and that represents the greatest diversity in landscapes of the three tiles covering the area. By employing this study area the results, conclusions and contributions made by the present research can be more easily applied to the current remote sensing-based mapping and modeling efforts of the FMFGBRP. At this time these remote sensing efforts do not involve the type of multi-temporal NDVI data set as that which forms the basis of the present investigation, nor the type of phenological information that can be extracted from such data sets, despite the demonstrable reliance of grizzly bear movement and activity on seasonally-available high-energy food resources (Munro *et al*. 2005, McDermid *et al*. 2006). Nevertheless, future collaborations with members of the FMFGBRP involving the

incorporation of this information into current remote sensing efforts provided the motive for selecting the location and extent of the current study area (Figure 3-2).

The study area comprises a variety of physiographic and biological landscapes, from grassland and wooded landscapes in the east to mountainous terrain in the west. For the purposes of governmental resource management and planning, Achuff and Wallis (1977) and later Achuff (1992), classified the diversity of Alberta's environments into a number of natural regions and subregions, five of which fall within the study area: Grassland, Parkland, Boreal, Foothills and Rocky Mountain regions (Figure 3-3). This hierarchy forms the basis of landscape organization in the current research. Representing the highest level in Achuff's (1992) hierarchy, these natural regions share common but broad vegetation, physiographic, and soil patterns (Achuff 1992), while more complex patterns of soil, vegetation and climate further divide these into natural subregions. Table 3-1 provides a summary of the major vegetative and physiographic characteristics of the natural regions and subregions found in the study area. Both Banff and Jasper National Parks are located in the study area, but increasing resource extraction and other economic activities characterize much of the remaining landscape (Figure 3-2).

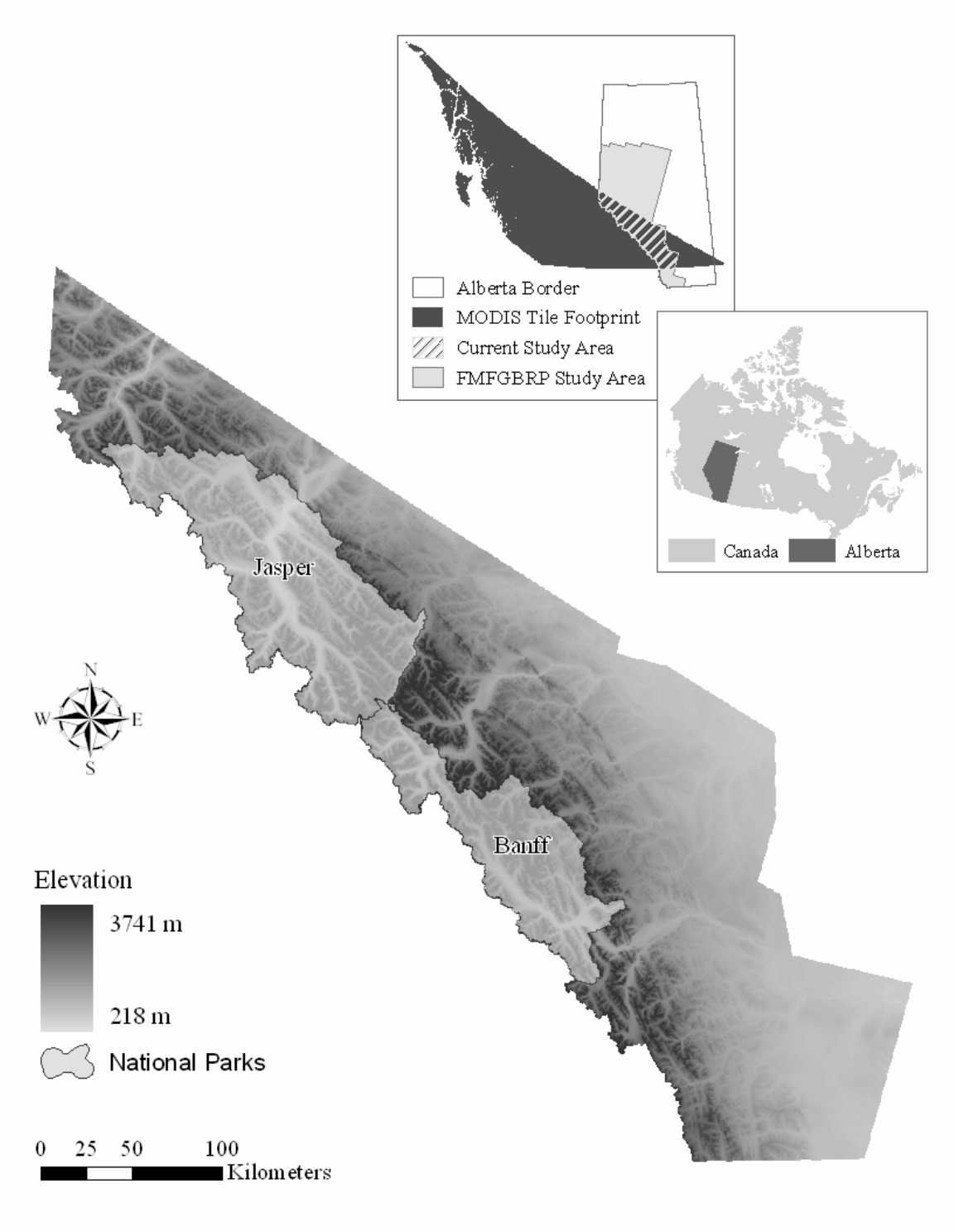

Figure 3-2: Map showing the location, context and extent of the study area within west-central Alberta, Canada.

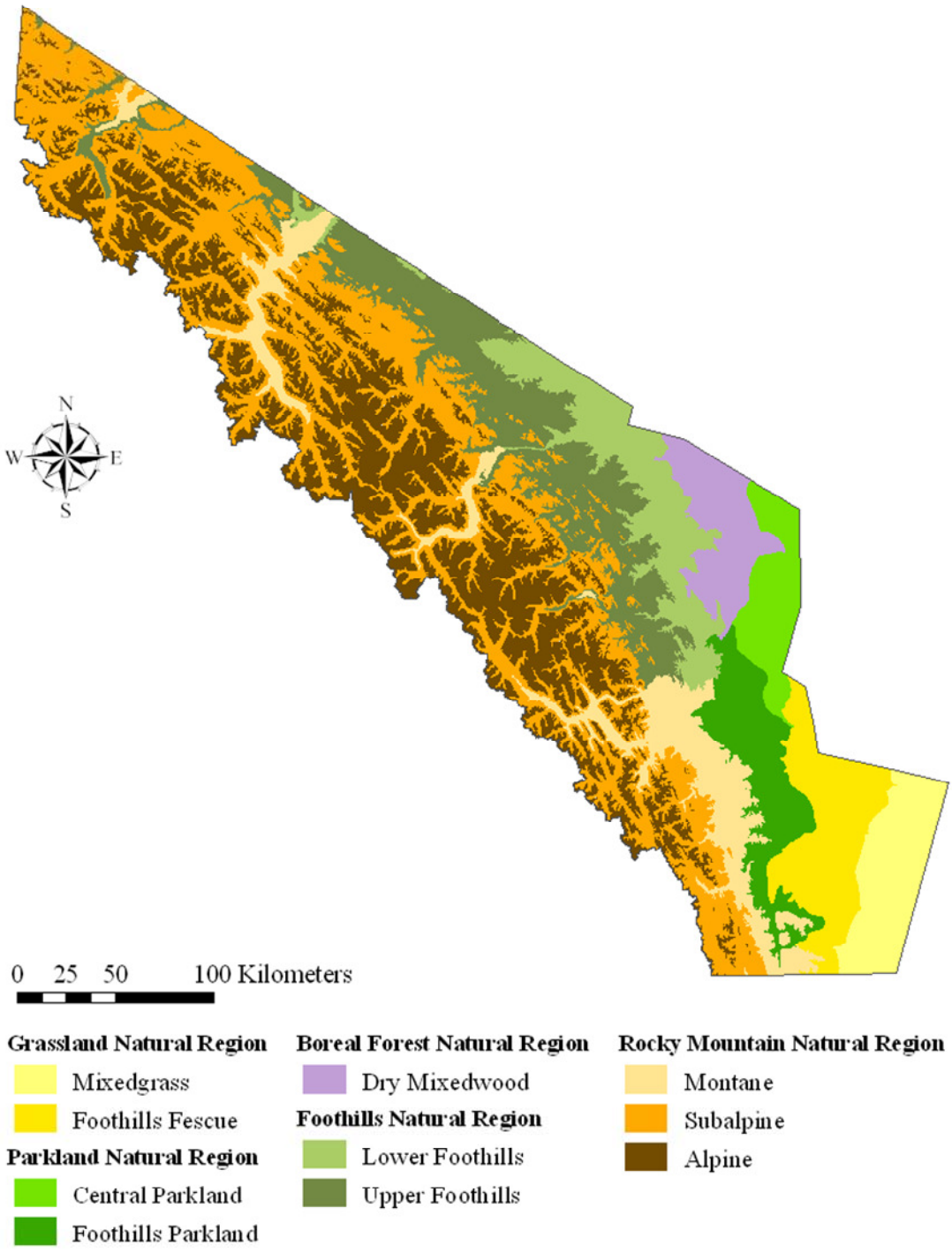

Figure 3-3: The natural regions and subregions comprising the study area. Data source: Government of Alberta (2005).

Table 3-1: Characteristics of the natural regions and subregions present in the study area. Information compiled after Achuff, 1992.

Region	Subregion	Geology and Landforms	Vegetation
Grassland	Mixedgrass	Hummocky ground moraine and fine glaciolacustrine deposits, subdued terrain, with minor uplands	Spear grass, wheat grass, porcupine grass, snowberry, silverberry, prickly rose, narrowleaf cottonwood
	Foothills Fescue	Flat to gently rolling terrain, consisting of moraine and glaciolacustrine deposits	Rough fescue, Idaho fescue, oatgrass, sticky geranium and prairie corcus, balsam-root, narrowleaf cottonwood
Parkland	Central Parkland	Hummocky ground moraine and fine glaciolacustrine deposits	Aspen poplar, balsam poplar, snowberry, saskatoon, bunchberry, red-osier dogwood, willow, alder, rough fescue
	Foothills Parkland	Hummocky ground moraine, outwash deposits, and extensive river terraces	Aspen poplar, balsam poplar, snowberry, saskatoon, white meadowsweet, glacier lily, Bebb's willow
Boreal Forest	Dry Mixedwood	Low relief topography consisting of ground moraine and sandy outwash plain	Aspen poplar, balsam poplar, white spruce, balsam fir, jack pine, black spruce, tamarack, cranberry, red-osier dogwood, feathermoss, bearberry, lichen, Labrador tea, peatmoss, sedge
Foothills	Lower Foothills	Rolling topography consisting of moraine deposits over folded bedrock; extensive organic deposits in valleys and wet depressions	White spruce, black spruce, londgepole pine, balsam fir, aspen poplar, paper birch, balsam poplar, buffaloberry, juniper Labrador tea, fireweed, dwarf birch, horsetail, willow, peat moss
	Upper Foothills	Strongly rolling topography with frequent bedrock outcrops; ground morain over bedrock with some colluvium on steep terrain	White spruce, black spruce, londgepole pine, buffaloberry, bunchberry, Labrador tea, fireweed, feathermoss, dwarf birch, peat moss
Rocky Mountain	Montane	Fluvial and glaciofluvial terraces and deposits in road river valleys	Douglas fir, limber pine, white spruce, aspen poplar, pine grass, hairy wild rye, bearberry, juniper, wheatgrass, Idaho fescue
	Subalpine	Rugged terrain comprising morainal and colluvial deposits overlying Rocky Mountain strata.	Lodgepole pine, Englemann spruce, subalpine fir, whitebark pine, buffaloberry, aster, hairy wild rye, junipers, grouseberry, false azalea, huckleberry, bunchberry, arroweaf groundsel, willow
	Alpine	Extremely rugged terrain containing bedrock, glaciers, snowfields and colluvium; rock glaciers and other permafrost features also present	Back alpine sedge, heathers, grouseberry, mountain avens, willow, moss, lichens

The Rocky Mountain region covers the majority of the study area, (approximately 58% of the total area), stretching from the northwest to the southeast along its western edge (Figure 3-3). Comprised largely of up-thrust, faulted sedimentary strata overlain by glacial deposits, this rocky and rugged terrain ranges from less than 1000 m to over 3700 m above sea level. Trending north-west to south-east, the major valleys encompass considerable variety in soils, vegetation, climate and biological characteristics. These change with elevation, shifting from the open forests and grasslands of the Montane subregion in the lower elevations that are home to many wildlife species such as mule deer, bighorn sheep, elk and wolf, to the conifer-dominated mesic forests in the Subalpine subregion, to the thin soils and rocky, treeless terrain of the highest elevations in the Alpine subregion, in which woodland caribou, bighorn sheep, mountain goat, grizzly bear and ptarmigan are seasonal inhabitants (Achuff 1992).

Covering the second largest portion (approximately 20%) of the study area (Figure 3-3), the Foothills natural region encompasses rolling topography underlain by folded sandstone and shale bedrock. The Upper and Lower Foothills, two subregions found to the east of the Rocky Mountain region, show distinct vegetation, physiography and wildlife communities. The Upper Foothills consist of strongly rolling topography along the western edge of the study area, dominated by mostly coniferous forests that are generally home to elk, black bears, grizzly bears, pine siskin and thrushes (Achuff 1992). The Lower Foothills, in the eastern portion of the region, also comprise rolling topography but are more varied in terms of vegetation communities, which include mixed forest and low-lying wetlands, with a diversity of wide-ranging wildlife such as elk, moose, spruce grouse, and purple finch (Achuff 1992).

The Grassland natural region occupies approximately 10% of the study area in the south-eastern corner (Figure 3-3). Consisting of relatively flat to rolling terrain, these grasslands are covered by glaciolacustrine deposits that support various grasses and forbs, along with narrowleaf cottonwood tree species, where native vegetation remains. Much of the region is now occupied by agricultural activity, however, which flourishes in the rich soils and warm, dry climate (Achuff 1992). A variety of birds, including grouse, sparrows, and thrashers, as well as rabbits, ground squirrels and deer species are found

across this region. The Mixedgrass and Foothills Fescue subregions both fall within the study area, and differ in terms of climate, soil and vegetation – the latter shows less extensive riparian cottonwood stands than the former, and much higher elevations, up to 1400m (Achuff 1992).

Covering approximately 8% of the total area, the Parkland natural region runs down the south-central portion of the study area, where two subregions stretch along a north-south gradient (Figure 3-3). In the Central Parkland subregion to the north, comprising glacial hummocky and flat outwash terrains, the naturally-occurring aspen stands separated by open grassy areas have been predominantly replaced by agricultural developments (Achuff 1992). White-tailed deer, ground squirrels, porcupine and snowshoe hare along with various birds and amphibians can still be found in the scattered pockets of wilderness that remain. South of the Central Parkland, the Foothills Parkland natural subregion is physiographically similar but is differentiated by a variety of distinct local grasses and flowers where native vegetation is still found (Achuff 1992).

Located toward the east-central portion of the study area (Figure 3-3) and comprising less than 4% of this area, the Boreal Forest natural region consists of forested lowland plains and wetland areas. Of the six natural subregions comprising this diverse region, only one is present in the study area: the Dry Mixedwood subregion. This subregion is covered by predominantly broadleaf upland forests interspersed with wetland, and less common coniferous stands (Achuff 1992). Black bear, moose, and beaver are typically found here, though animal species are not as diverse here as they are in subregions found further northward.

The study area shows considerable diversity in geology, topography, vegetation, climate and wildlife across the five regions and ten subregions that comprise it. With such an assortment of ecosystems and environments a categorization of the landscape into relatively homogeneous spatial units, such as that provided by Achuff (1992), was necessary for the proper treatment of the study area in the current research. Therefore, these natural regions and subregions provided the basis for landscape organization for this project.

3.2 MODIS Data Acquisition and Pre-processing

A set of MODIS MOD13Q1 (Level 3, Global, Version 004) 16-day, 250m, NDVI images from the Terra satellite covering the study area from January 1, 2003 to December 31, 2005, provided the satellite data used in the current research (Table 3-2). Each year comprises 23 16-day composites, for a total of 69 images over the three years. These data are distributed free of charge by the Land Processes Distributed Active Archive Center (LP DAAC) located at the U.S. Geological Survey Center for Earth Resources Observation and Science. The imagery was accessed through the EOS Data Gateway website (http://edcimswww.cr.usgs.gov/pub/imswelcome/), and was provided in NASA's Hierarchical Data Format-Earth Observing System (HDF-EOS) format in an integerized sinusoidal grid projection (Huete *et al*. 1999, LP DAAC 2005). Each MOD13Q1 image contains 11 bands: NDVI, NDVI Quality Assurance Science Data Set (QASDS), EVI, EVI QASDS, red, near infrared, blue, and middle infrared reflectance, average view and sun zenith angles, and average relative azimuth angle (Table 3-2; LP DAAC 2005). The MODIS re-projection tool (version 3.0), downloaded from the LPDAAC website, was used to extract the NDVI and NDVI Quality bands, re-project these data by nearest-neighbour resampling into the NAD83, UTM Zone 11N projection, and convert them into a more useable format (e.g. TIFF files). The multi-temporal data sets were then clipped to the study area using the Research Systems, Incorporated ENVI 4.2 software package. TheNDVI imagery was provided as scaled values (i.e. multiplied by 10,000) and was left in this scaled format for portions of the processing and analysis in order to minimize data storage requirements (i.e. it was left in a 16-bit signed integer format, rather than the floating point format required by unscaled NDVI values between -1 and 1). Data stacks of the 69 NDVI and QASDS images covering the three years were clipped to the extent of the study area using ENVI 4.2. These clipped data sets formed the basis of the subsequent model building.

Table 3-2: Characteristics of the MODIS data product used in the current research. Information compiled from LP DAAC (2005).

Satellite:	Terra
Sensor:	Moderate Resolution Imaging Spetroradiometer (MODIS)
Coverage:	Global
Data Set Name:	MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN
Data Format:	HDF-EOS (Hierarchical Data Format - Earth Observing System)
Projection:	Sinusoidal
Granule Shortname:	MOD13Q1
Area of Tile:	~ 10^0 Latitude by 10^0 Longitude
Resolution:	250 m
Composite Period:	16 days
Version:	004
Science Data Sets:	NDVI*, EVI, NDVI Quality*, EVI Quality, Red Reflectance, NIR Reflectance, Blue Reflectance, MIR Reflectance, Average View Zenith Angle, Average Sun Zenith Angle, Average Relative Azimuth Angle

**Used in current study.*

3.3 Constructing the Model Environment

As explained above, the cost and difficulties associated with incorporating ground truth into a real-world study of noise reduction in NDVI time series necessitated the use of a model environment for the current empirical comparison of selected noise reduction techniques. This model environment consisted of six modeled NDVI time series representing various land cover types from within the study area, to which three different levels of noise were introduced, representing slight, moderate and high levels of noise. In order to reproduce the real-world effects of noise, the introduced noise was derived from actual noisy pixels extracted from the data set. The six original model time series and the 18 model time series with introduced noise, to which each of the selected noise reduction techniques were applied, formed the basis of current investigation. Thus, the performance of each selected technique could be evaluated across different land covers

and intensities of noise. Although the model environment was idealized and could only reflect the reality of noisy NDVI time series to a point, it is expected that basing model construction on actual NDVI time series would serve to reflect real-world conditions.

3.3.1 Model Time Series Construction

Figure 3-4 presents the six model time series that were constructed from the multi-temporal MODIS NDVI data set covering the study area for the years 2003 through 2005. Each reflects a different land cover type from within the study area and represents a combination of the highest quality, most noise-free pixels from a particular region or subregion (Figure 3-3). The construction of the six model NDVI time series, designed to reflect the various land cover types found in the study area and their respective phenological cycles, was based on the assumption that each natural subregion within the study area represents a largely homogeneous land cover type, and that each land cover type displays a distinctive yearly phenological cycle that would be reflected in a time series of NDVI values. The terms 'land cover' and 'land cover type' are used in the present work to denote a large-area, homogenous eco- or bioregion containing a generally consistent type of vegetated surface (e.g. mixed boreal forest) rather than individual ground features, and must be interpreted as such. In order to construct model time series for each land cover type, individual, separate data sets covering the extent of each subregion were extracted from the original data set using ENVI 4.2. These individual data sets formed the basis of subsequent model building.

Once the individual subregion data sets had been extracted, the highest quality pixels were selected from each. It was assumed that in using pixels with the highest quality data for the three years, as identified through quality assurance (QA) flags provided with the MOD13Q1 data, the most accurate real-world time series could be modeled for each subregion. An average time series calculated from these selected pixels formed the foundation of each of the model time series. In this way, the model time series were based on actual NDVI data of the highest quality.

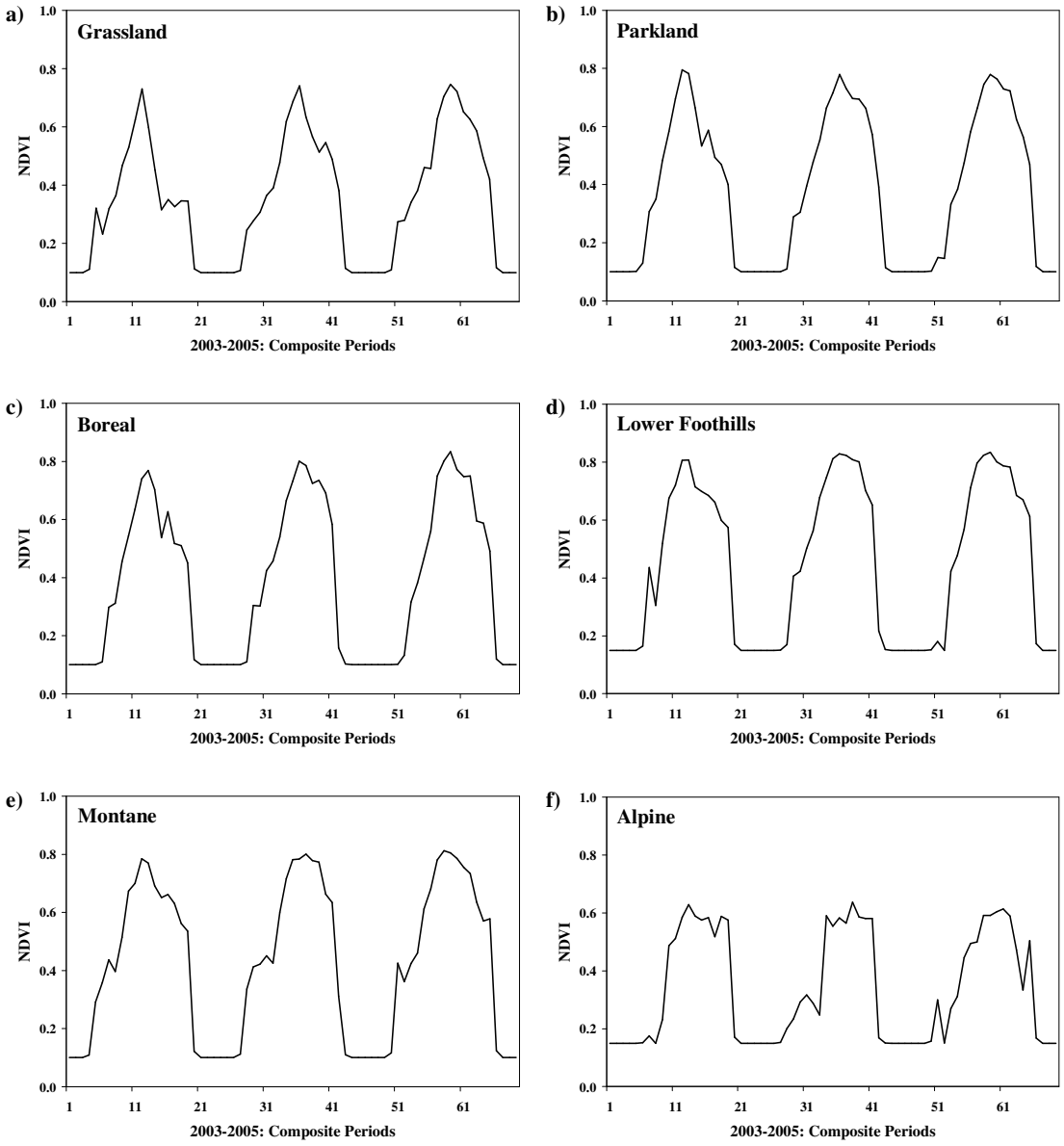

Figure 3-4: Final model NDVI time series for the a) Grassland, b) Parkland, c) Boreal, d) Lower Foothills, e) Montane and f) Alpine land cover types.

Selecting the best pixels from each subregion involved the use of the vegetation index (VI) usefulness index, which is one of several per-pixel data sets that are available with the NDVI QASDS. Simple code, written in Interactive Data Language (IDL) and applied in the Research Systems, Incorporated IDL 6.2 software, provided a means of extracting the usefulness index from the QASDS. As a quality rating for the NDVI estimates provided in the MOD13Q1 data set, the usefulness index ranges from 'perfect' (0) to 'unusable' (15) and is based on the combined affects of atmospheric, aerosol,

cloud, snow, and shadow contamination present in the data (LP DAAC 2005). Identifying the least noisy (i.e. cleanest) pixels involved the selection of those pixels with the greatest number of NDVI values possessing usefulness indices of 'good' or higher quality (i.e. a usefulness index of 2 or less). This process required the definition of a minimum threshold to distinguish those pixels with sufficient number of 'good' NDVI estimates to be used in constructing a model time series for each subregion, from the rest. Adjustment of this threshold was required for each of the subregions because the greatest numbers of 'good' NDVI estimates varied between subregions. The threshold varied from 26 'good' quality NDVI estimates for pixels in the Alpine subregion, to 42 'good' estimates for pixels in the Grassland subregion, but in every case was adjusted so that a minimum of 100 pixels would be selected for model construction. A simple IDL code provided the means of setting this threshold and selecting the highest-quality pixels for each subregion. Once selected, the NDVI time series for all the best pixels for each subregion were averaged on a date-by-date basis to produce the initial set of modeled NDVI time series.

The initial model time series required further processing before they could reflect ideal conditions, however. For instance, despite their basis on an average of the best pixels in a subregion, some of these time series were observed to contain negative NDVI values. Because such values represent an absence of vegetation, and because it was assumed that each of these ideal time series represented vegetated surfaces, all negative NDVI values were assumed to represent data errors and were set to zero (Myneni *et al.* 1995). An illustration of this is found in Figure 3-5.

In addition, winter NDVI values for the months of November through February demonstrated undesirable fluctuations during a period which, based on general knowledge of regional climate and phenological development, should generally show constant NDVI representing winter vegetation. In other words, spurious spikes and drops occurring during a period when photosynthesis should be at a minimum (in evergreen species) or altogether dormant, does not correspond to an ideal time series. An inspection of the VI usefulness index flags for the selected pixels revealed that these fluctuations were found to correspond to the lowest quality NDVI estimates in the time

series, with the highest quality estimates occurring between June and August. In light of this lower quality and the knowledge of local climatic and phenological conditions, all winter NDVI values in the model time series representing composite periods from November through February were reset to an NDVI of 0.1, which corresponds to and NDVI for bare ground (DeFries and Townshend 1994). The Alpine and Lower Foothills subregion model time series were an exception; for these a higher average minimum NDVI was observed in the time series, and a winter NDVI of 0.15 was therefore chosen instead (Figure 3-4). In this way the undesirable fluctuations in winter NDVI were eliminated and the model time series approximated ideal phenological profiles more closely.

One final adjustment was made to the model time series before they were considered to represent ideal conditions. This involved a softening of the transition between winter dormancy and the growing season (i.e. at the onset of spring and at the end of autumnal senescence) because a very sudden and angular transition was observed in all model time series. The suddenness of these transitions was likely a result of the composite period length of 16 days, a temporal resolution which is perhaps not optimal for capturing the relatively rapid transitions to snowmelt and green-up, or to winter dormancy in the study area. These awkward transitions do not represent ideal conditions. It is commonly assumed that vegetation growth and development is relatively gradual and progresses smoothly over time without sudden sharp increases or decreases, and that NDVI time series should reflect this (van Dijk *et al*. 1987, Ma and Veroustraete 2005, Wang *et al*. 2005b). On the basis of this assumption, the first and last NDVI composite periods for each season of winter NDVI values was increased by 5% of the NDVI range between dates before and after it. In this way the transition between winter dormancy and spring green-up was smoothed slightly and final, idealized model NDVI time series produced (Figure 3-4). These final NDVI time series formed the basis of the model environment in which the subsequent empirical comparison was undertaken (Figure 3-3).

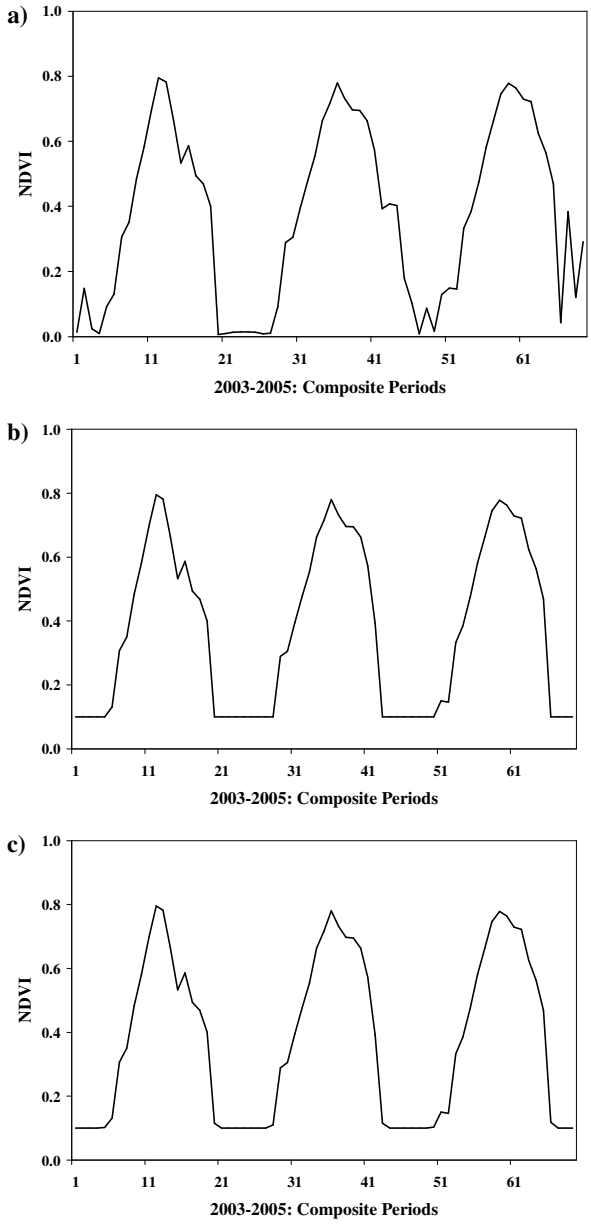

Figure 3-5: An example of the steps used in model construction, showing: a) the averaged NDVI time series for the Parkland region, b) the setting of all winter NDVI values to 0.1, and c) the final Parkland model NDVI time series with adjusted transitions at the beginnings and ends of the growing seasons.

No processing was performed on non-winter values in any of the model time series because these dates represented the highest quality original values, and therefore presumably the most accurate NDVI values in the data set. All but two of the subregion model time series, the Upper Foothills and the Subalpine produced relatively smooth model time series. Because of the level of noise still present in these two time series,

they were dropped from the subsequent analysis. In addition, because the profiles for the two Grassland subregions were extremely similar, they were combined into one Grassland model time series in order to reduce redundancy in the analysis. A similar observation was made for the two Parkland subregions, and the two subregions were therefore also combined to produce one Parkland model time series for the analysis.

3.3.2 Introducing Noise

Noise was introduced to the model time series by randomly selecting 10%, 40% and 70% of the dates from each modeled time series and replacing those ideal NDVI values with real NDVI values from a noisy pixel locatied within the appropriate subregion. As explained previously, these three levels of introduced noise were chosen to represent slightly, moderately and very noisy time series. This process involved first selecting the noisiest pixel from the subregion or region represented by a model time series, identified as the pixel showing the highest number of poor NDVI values. The MODIS VI usefulness index was used to detect poor NDVI values, where indices of 'average' (8) or worse (>8) were identified using IDL code. The pixel with the highest number of average- or worse-quality values represented the worst pixel in the region or subregion, and provided the source of noise for the appropriate model time series. The three sets of randomly chosen dates were then replaced with the corresponding value for that date from the noisy pixel. Realistic noise was thus introduced at random at three different levels to each of the model time series. An illustration of this is provided in Figure 3-6. The 18 noisy time series resulting from this process are provided in Appendix A.

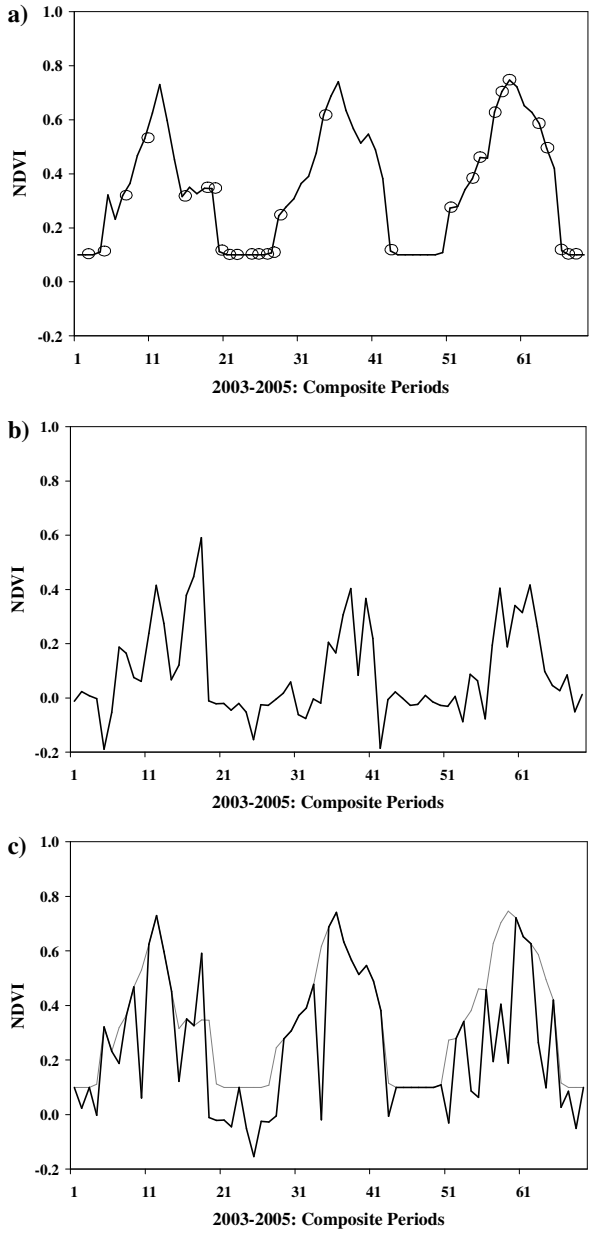

Figure 3-6: An example of noise introduction to the Grasslands model NDVI time series, showing: a) the model time series with 40% of the dates selected (circles), b) the noisy pixel time series selected from the region, and c) the model NDVI time series with the appropriate dates replaced by corresponding values from the noisy pixel (the original model time series is shown in grey).

3.4 Noise Reduction Technique Selection and Implementation

3.4.1 Selecting Candidate Noise Reduction Techniques

The literature review presented in Chapter Two provided the basis for selecting a series of candidate noise reduction techniques for empirical comparison. It was not feasible within the scope of the current project to compare all the methods outlined in the previous chapter; therefore, a more practical approach was necessary. This involved choosing techniques that not only demonstrated successful application in the literature, but were likely to be easily accessible and applicable to a variety of future research objectives, and would thus generally represent the most potentially useful methods for the current research and study area. Of the techniques outlined in Chapter Two, two function-fitting methods and four filtering procedures were chosen for comparison. They represent a range of approaches to the reduction of noise in NDVI time series, and are presented in Table 3-3.

Table 3-3: The six noise reduction techniques selected for empirical comparison.

Category	Technique	Reference(s)
Function-Fitting	Asymmetrical Gaussian Function-Fitting	Jönsson *et al* . (2002)
	Double Logistic Function-Fitting	Beck *et al* . (2006)
Filters	Savitzky-Golay Filter	Chen *et al.* (2004)
	4253H-Twice Filter	Velleman (1980)
	Mean-Value Iteration Filter	Ma *et al* . (2005)
	Combined Autoregressive Filter (ARMD3-ARMA5)	Hall-Beyer *et al* . (unpublished)

Two function-fitting techniques were selected for empirical comparison: Jönsson *et al*.'s (2002) Asymmetrical Gaussian function-fitting and Beck *et al*.'s (2006, 2007) Double Logistic function-fitting. Both of these algorithms showed successful application in the literature, and were easily implemented using Jönsson and Eklundh's (2002, 2004) TIMESAT program, available from the authors. Jönsson and Eklundh (2002) demonstrated the superior performance of the Asymmetric Gaussian method over both a Fourier-based technique and the BISE algorithm, while Beck *et al*. (2006) presented

similar results for the Double Logistic method when compared with Fourier analysis, and was shown to better account for snow and high sun angle effects at higher latitudes. Both the Asymmetric Gaussian and Double Logistic techniques were found to perform similarly when applied to a noisy NDVI time series (Beck *et al*. 2006). However, because both techniques demonstrated superior performance and successful application, and are easily accessible, it was considered desirable to include both in the selection of candidate techniques for empirical comparison.

Alternative function-fitting techniques to the Asymmetric Gaussian and Double Logistic either did not perform as well when tested in the literature or were less practical in their application to noise reduction for NDVI time series. For instance, despite the demonstrated popularity of Fourier transforms for noise reduction in NDVI time series, several distinct disadvantages are recognized. In particular, these methods require equal spacing of values in time (Jönsson and Eklundh 2002), can be unsuccessful over time series containing considerable noise and error (Chen *et al*. 2004) or where time series are not symmetrical and regular (Roerink *et al*. 2000), and can generate spurious oscillations in the fitted curve (Ma and Veroustraete 2005) or curves far removed from the original data values (van Dijk *et al*. 1987). The HANTS algorithm, though designed to deal with some of these issues, requires the user to subjectively set five control parameters on the basis of personal experience and trial and error (Roerink *et al*. 2000), unlike either of the two selected function-fitting approaches. Although similar in nature to the Beck *et al*.'s (2006, 2007) Double Logistic function-fitting, Zhang *et al*.'s (2003) Piecewise Logistic function is not aimed at noise reduction, but rather, at the extraction of phenological information. That is, the algorithm assumes that only minimal noise exists. In addition, the technique does not account for the biased effects of cloud, aerosol and snow contamination, or possible bidirectional effects.

Neither Zhang *et al*.'s (2003) method, nor the spline and wavelet-based methods described in Chapter Two were considered as easily applicable or advantageous as the two selected function-fitting techniques. The spline-based method described in Bradley *et al*. (2007) appears promising, however, the details of its implementation were only very recently available (Hermance *et al*. 2007). In addition, both spline and wavelet-

based methods are little used in the literature, not well described (e.g. Wang *et al*. 2005a, 2005b), and could both present mathematically complex techniques requiring parameter manipulation by experienced users.

Four filtering techniques were selected for empirical comparison because of their flexibility, simplicity and ease of use. These are: the Savitzky-Golay filter (Chen *et al*. 2004), the 4253H-Twice filter (Velleman 1980), the MVI filter (Ma and Veroustraete 2005), and the ARMD3-ARMA5 filter (Filipova-Racheva and Hall-Beyer, unpublished).

The Savitzky-Golay filter presents a flexible NDVI time series noise reduction technique that can be applied to data at a variety of spatial and temporal scales, and has been shown to outperform both Fourier-based methods and the BISE algorithm (Chen *et al*. 2004). The filter does not produce spurious oscillations as in the case of Fourier analysis, nor does it require the setting of pixel-by-pixel thresholds and a sliding period, as is necessary for the BISE method (Chen *et al*. 2004). Applicable in Jönsson and Eklundh's (2004) TIMESAT program, this filter is easily accessible and presents a practical method for reducing noise in NDVI time series.

Although only applied to NDVI time series by van Dijk *et al*. (1987), Velleman's (1980) 4253H-Twice filter was shown to provide the most effective noise reduction over running weighted mean, simple running median filters, and Tukey's (1977) 3RSSH-Twice filter. The former did not result in profiles very far removed from the original trends in the time series, as did the moving mean filter, nor were the resulting time series as rough as the simple running median filters (van Dijk *et al*. 1987). By running a series of median windows of varying sizes, followed by a 'reroughing' procedure, the 4253H-Twice filter minimizes these concerns.

Also demonstrating more efficient and effective noise reduction for NDVI time series than the Fourier analysis and the BISE algorithm (Ma and Veroustraete 2005), the MVI filter was chosen as the third filtering technique for empirical comparison (Table 3-3). Although it also only appears once in the literature, the MVI filter shows promise as a noise reduction strategy. Not only is it straightforward, uncomplicated, and flexible, but its iterative nature provides a means of reducing noise through continuous re-evaluation.

The final candidate noise reduction technique, Filipova-Racheva and Hall-Beyer's (unpublished) ARMD3-ARMA5 filter also provides a simple and easily-applied method for noise reduction (Table 3-3). The authors demonstrated the superior performance of this autoregressive technique over other, alternative autoregressive filters. Empirical comparison provides a valuable framework in which the applicability of such a little-used technique can be quantitatively assessed in terms of its comparative performance to other more well-known techniques, as in the case of the MVI filter. The simplicity, ease of use and demonstrated potential of the Savitzky-Golay, 4253H-Twice, MVI and ARMD3-ARMA5 filters support the value of testing their comparative performances in NDVI time series noise reduction.

Simple averaging filters (e.g. Chen *et al*. 2000, Xiao *et al*. 2002) and atmospheric corrections (Rahman and Dedieu 1994, Kobayashi and Dye 2005), while useful, were not selected for the current comparison, however. The former, as explained previously, can result in time series that deviate significantly from original profile trends (van Dijk *et al*. 1987), while the latter does not account for bidirectional effects, which are not insignificant in MODIS NDVI time series (van Leeuwen *et al*. 1999). The BISE algorithm (Viovy *et al*. 1992), the ABC3 technique (Cihlar *et al*. 1997, 2004), and Kang *et al*.'s (2005) cloud interpolation scheme were also not selected for empirical comparison. Not only has the inferiority of the BISE method to the selected techniques been demonstrated (Jönsson and Eklundh 2002, Chen *et al*. 2004, Ma and Veroustraete 2005), but the need for subjectively-set per-pixel thresholds and a sliding period, renders this technique less easily applied. In the case of Cihlar *et al*.'s (1997, 2004) ABC3 procedure , the need for extensive ancillary data sets that are not always easily accessible and its complex implementation make this method inappropriate for wide-ranging applications. Similarly, the reliance of Kang *et al*.'s (2005) cloudy pixel interpolation scheme on regular, relatively clean (i.e. non-cloudy) pixels and dates would not be effective in regions of prevalent cloud cover, and at particular cloudy times of year for most global land covers. Bidirectional effects are also unaccounted-for in this technique.

3.4.2 Candidate Noise Reduction Techniques: Description and Implementation

3.4.2.1 Asymmetrical Gaussian Function-Fitting

The Asymmetrical Gaussian function-fitting technique described by Jönsson and Eklundh (2002, 2004) fits local, nonlinear model functions at intervals around local maxima and minima in a time series of NDVI values. These local functions are merged to build global functions that describe NDVI over full seasonal cycles. The method is flexible in that it can be applied to time series at varying temporal resolutions (e.g. daily, bi-weekly, monthly values), and with scaled or unscaled NDVI values. The procedure was implemented using Jönsson and Eklundh's (2004) TIMESAT 2.3 program run in MATLAB 7.1 software. The following description is taken from the TIMESAT 2.3 User Guide (Jönsson and Eklundh 2006), but some details are added from Jönsson and Eklundh (2002, 2004) for clarity.

First, the procedure uses the following model function to determine local maxima (Jönsson and Eklundh 2006):

$$f(t) = x_1 + x_2(wt) + x_3\cos(wt) + x_4\sin(2wt) + x_4\cos(2wt) \qquad (3\text{-}1)$$

where t is the independent variable, time, $x_1 \ldots x_4$ are parameters, w is calculated as $\frac{6\pi}{N}$, and N represented the number of data values in the time series. Local model functions are fit at intervals around these local maxima and minimna, and have the general form (Jönsson and Eklundh 2002):

$$f(t) \equiv f(t; c_1, c_2, a_1, \ldots, a_5 = c_1 + c_2 g(t; a_1, \ldots, a_5) \qquad (3\text{-}2)$$

where c_1 and c_2 are parameters to determine function base level and amplitude, and a_1 through a_5 are parameters to determine the position of the local maxima or minima with respect to t, the width of the right function half, the flatness of the right function half, the width of the left function half, and the flatness of the left function half, respectively. In addition, the following is a Gaussian-type function (Jönsson and Eklundh 2006):

$$g(t;a_1,\ldots,a_5)=\begin{cases}\exp\left[-\left(\dfrac{t-a_1}{a_2}\right)^{a_3}\right], & \text{if } t>a_1\\ \exp\left[-\left(\dfrac{a_1-t}{a_4}\right)^{a_5}\right], & \text{if } t<a_1\end{cases} \tag{3-3}$$

Parameters a_1 through a_5 are restricted to within a reasonable range to ensure smooth model function shapes for the resulting curves. The following equation is used to determine the parameters listed above (Jönsson and Eklundh 2002):

$$x^2=\sum_{i=n_1}^{n_2}\left[\frac{f(t_i;c_1,c_2,a_1,\ldots,a_5)-I_i}{\sigma_i}\right]^2 \tag{3-4}$$

where x^2 is a merit function, i is a data point, between $n_1,\ldots,n_2$, t_i is time at point i, I_i is the NDVI value at point i, and σ_i is the known uncertainty of point i, or is set to 1. In order to account for the negative bias produced by atmospheric affects and adopt the upper envelope of values in the fit, this equation is applied twice. Data points found below the model function fit after the first iteration, are weighted with higher uncertainty in σ_i during the second iteration (Jönsson and Eklundh 2002).

The final step in the process is to merge these local functions for the left minimum, central maximum and right minimum, denoted as $f_L(t)$, $f_R(t)$, and $f_C(t)$, into a global function $F(t)$ that models the full NDVI variation (t_L,t_R), using the following (Jönsson and Eklundh 2006):

$$F(t)=\begin{cases}\alpha(t)f_L(t)+[1-\alpha(t)]f_C(t), & t_L<t<t_C\\ \beta(t)f_C(t)+[1-\beta(t)]f_R(t), & t_C<t<t_R\end{cases} \tag{3-5}$$

where $\alpha(t)$ is , and $\beta(t)$ are cutoff functions around $\dfrac{(t_L+t_C)}{2}$ and $\dfrac{(t_C+t_R)}{2}$, respectively.

The TIMESAT 2.3 program is available online through the authors (Jönsson and Eklundh 2006), and evaluates NDVI time series on a pixel by pixel basis. The model time series with 10%, 40% and 70% introduced noise were input to the program as individual pixels. See Jönsson and Eklundh (2002, 2004, 2006) for a more detailed description of the TIMESAT program.

3.4.2.2 Double Logistic Function-Fitting

Also implemented using TIMESAT 2.3, the Double Logistic function-fitting procedure employs Equation (3-1) to determine local maxima. This technique models a one-year NDVI time series as a function of time using the following basis function (Jönsson and Eklundh 2006):

$$f(t;d_1,\ldots,d_4)=\left(\frac{1}{1+\exp\left(\frac{x_1-t}{x_2}\right)}\right)-\left(\frac{1}{1+\exp\left(\frac{x_3-t}{x_4}\right)}\right) \tag{3-6}$$

where d_1 is the position of the left inflection point, d_2 is the rate of change at d_1, d_3 is the position of the right inflection point, and d_4 is the rate of change at d_3. Equation (3-4) is used to derive these parameters as it is for the Asymmetrical Gaussian function-fitting (Jönsson and Eklundh 2006), while the biased effects of cloud and aerosol contamination are accounted for using weights. A first fitting in the Double Logistic procedure is used to derive the weighting scheme for growing season observations based on residual analysis, whereby lower observations are given less weight, and to re-estimate all parameters used in the equation (Beck *et al*. 2006, 2007). The final result emphasizes higher NDVI values, assumed to reflect the most accurate estimates. This technique assumes that all NDVI time series should follow a smoothly-varying, pre-defined curve and does not account for sudden drops in NDVI during the growing season. In other words, the effects of sudden droughts, fires or other disturbances are ignored during the function-fitting (Beck *et al*. 2006).

3.4.2.3 The Savitzky-Golay Filter

Based on the simple least-squares-fit moving window filter described by Savitzky and Golay (1964), the adjusted Savitzky-Golay filter presented by Chen *et al*. (2004) applies a weighted moving average filter to an NDVI time series, with the weighting given as a polynomial of a particular degree. A polynomial least-squares fit is applied within the filter window by the weight coefficients. As in the case of the two noise reduction techniques described above, the Savtizky-Golay filter was the third and final technique implemented using Jönsson and Eklundh's (2004) TIMESAT 2.3 program.

The filtering procedure can be generally described by the equation (Chen *et al.* 2004, Jönsson and Eklundh 2006):

$$NDVI*_j = \frac{\sum_{i=-m}^{i=m} C_i NDVI_{j+i}}{N} \quad (3\text{-}7)$$

where $NDVI*_j$ is the new filtered NDVI for data value j, $NDVI_{j+i}$ is the original NDVI for data value j, N is the number of NDVI values in the time series, C_i is the coefficient for the i^{th} data value of the filter window, and m is the half-width of the filter window. In the TIMESAT program, a quadratic polynomial is fit to all points in the moving window, replacing the value at each data point with that of the polynomial (Jönsson and Eklundh 2006). Multiple steps are used, as in the two previous techniques, to adapt the procedure to capture the upper envelope of the NDVI time series, accounting for the biased effects of cloud and aerosol contamination. Three filtering iterations were performed with window sizes of five, six and seven composite periods, respectively. These were chosen on the basis of observations, taken from a visual inspection of the data and the resulting filtered time series using a graphical user interface, provided with the TIMESAT program. See the TIMESAT User's Guide (Jönsson and Eklundh 2006) for additional detail.

3.4.2.4 The 4253H-Twice Filter

As one of several nonlinear time series smoothers presented by Velleman (1980), the 4253H-Twice filter applies a series of running median filters of varying window sizes to the data set. A running median filter can be generally represented by (Velleman 1980):

$$z_t = med\{y_{t-u}, \ldots, y_t, \cdots, y_{t+u}\} \quad (3\text{-}8)$$

where z_t is the new NDVI value at time t, $med\{\ldots\}$ is the median of the values listed in the parentheses, y_t is the original NDVI value at time t, and u is an integer representing half the width of the running window length. The 4253H-Twice first involves the application of running medians with window lengths of four, two, five and then three to the NDVI time series, followed by a running weighted mean with a window length of three and weights of 0.25, 0.50, and 0.25, respectively. The latter is referred to

as ‘Hanning’, hence the ‘H’ in the filter’s name (Velleman 1980, van Dijk *et al*. 1987). The residuals of the resultant time series undergo the same process of running medians and Hanning, and are then added to the initial result to produce the final filtered time series. This last step is referred to as ‘re-roughing’, and is designed to retain original patterns in the data set that have become too smooth after the initial median and mean filters are applied; it accounts for the ‘Twice’ in the filter name.

No additional treatment of the very first and last values in the time series was performed. Although this is done in other cases as the filter is not able to smooth these particular values because they are not preceded or succeeded by additional values (van Dijk *et al*. 1987), it was considered unnecessary here. Only two of the 69 values in the current data set will remained untouched by the 4253H-Twice filter, and as they do not fall during the growing season but in mid-winter, they were not considered to be of primary concern.

In order to implement the 4253H-Twice filter on the model time series, a program was written in IDL using the IDL 6.2 software provided with ENVI 4.2 by Research Systems, Inc. This program applied the series of running mean and median filters to each NDVI time series with introduced noise individually, according to the specifications found in Velleman (1980). The IDL code for the program is provided in Appendix B.

3.4.2.5 The Mean-Value Iteration Filter

The Mean Value Iteration (MVI) filter is based on the assumption that an NDVI value in a time series will closely approximate the average of those values directly before and after it in the series (Ma and Veroustraete 2006). The filter involves an iterative procedure in which each element in the time series is compared with the mean of the values immediately preceding and succeeding it. If the absolute difference between the two is lower than a specified threshold (i.e. here, 0.1 of the multi-year average NDVI) and is no larger than it is for other elements in the time series, it is retained; but if it exceeds the threshold or is a larger difference than previously calculated in the time series, the NDVI value of the element of interest is replaced with this mean. Computer code for the MVI filter was obtained from the authors, and a translated version was written in IDL using the original code as well as the descriptions provided in Ma and

Veroustraete (2006). Model NDVI time series were processed individually in the IDL version of the program. Appendix B provides the code for this program.

3.4.2.6 The ARMD3-ARMA5 Filter

The final candidate noise reduction technique was not taken from published literature, but nevertheless, presents an intriguing alternative to the methods described above. It combines two autoregressive moving window filters: an autoregressive moving median algorithm with a window size of three elements, followed by an autoregressive moving average algorithm with a five-element window (Filipova-Racheva and Hall-Beyer, unpublished). These function as general moving average and median filters, but rather than using only original NDVI time series values in their window calculations, they combine the new, filtered NDVI values that precede the central value of interest in the window, with the (as of yet) unfiltered NDVI values that follow in the series; hence the term 'autoregressive'. For instance, in the five-element autoregressive moving average filter, and average of the original central and two succeeding NDVI values along with the two, now filtered, preceding NDVI values replaced the central window value. In this way, the results from the processing of previous values are incorporated into the processing of following values. This is particularly advantageous for the capture of gradual temporal processes, such as vegetative growth, from noisy signals (Filipova-Racheva and Hall-Beyer, unpublished). As was the case for the previous two filters, this process was implemented using a short program written in IDL, generated using the descriptions given by Filipova-Racheva and Hall-Beyer (unpublished). Again, this code is provided in Appendix B.

3.5 Empirical Comparison: Root Mean Square Error

As described above, the empirical comparison of the six selected noise reduction techniques was divided into two main evaluations. These were designed to test both i) the ability of the techniques to minimize overall noise in NDVI time series, and ii) the ability of the techniques to preserve the integrity of the original NDVI signal for the subsequent extraction of phenological metrics. The first evaluation was a root mean square error (RMSE) analysis, described in the present section, and the second involved

the evaluation of several metrics and their derivation from NDVI time series, described in section 3.6 below. In each, the noise-reduced time series were tested against the original model and noisy, uncorrected time series.

3.5.1 Root Mean Square Error

RMSE provides a measure of the mean difference between observed and predicted values (Willmott 1982), which in this case, are represented by original model time series NDVI values, and the NDVI values of a time series resulting from the reduction of introduced noise, respectively. This analysis offers a means through which the ability of each of the candidate noise reduction techniques to return a noisy NDVI time series to its original state, can be tested. Through the RMSE analysis a comparison was made between the original model time series before noise introduction and the model time series after noise reduction by the various candidate techniques. RMSE was calculated using the following equation (Willmott 1982):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (P_i - O_i)^2}{N}} \tag{3-9}$$

where P_i is 'predicted' value (i.e. the noise-reduced NDVI value) for data point i, O_i is the 'observed' value (i.e. the original modeled NDVI value) for data point i, and N is the number of data points. RMSE is calculated in the units of the data values, so in the current analysis, RMSE results are presented as a value of NDVI, which is itself unitless. A lower RMSE indicates a smaller difference between the original and noise-reduced model time series, which signifies a superior ability by that particular noise reduction technique to minimize noise while retaining the original trends in the time series. In other words, there is less mean difference between the model time series and the noise-reduced time series. This analysis was performed using the Microsoft Excel 2003 software, for the 108 noise-reduced NDVI time series and the 18 noisy, unfiltered NDVI time series. Those time series comparisons showing the lowest RMSE results were assumed to demonstrate the best noise reduction performance, because the noise-reduced time series was closest to the original model time series. Noise in the time series

has been minimized. Processing of the RMSE results involved the calculation of basic descriptive statistics including minimum, maximum, mean, and standard deviation, for each of the noise-reduced and noisy NDVI time series. Further data analysis is described in section 3.7 below.

3.6 Empirical Comparison: NDVI Time Series Metrics

The following section describes the investigation of several NDVI time series metrics and their derivation from model, noisy and noise-reduced time series. Their comparison provided an assessment of how well each of the candidate noise reduction strategies preserved the shape and amplitude of the original NDVI signal. Each of the time series metrics involved in the present analysis and their respective implementations are described below.

3.6.1 Selected Time Series Metrics

In order to compare the effects of the six candidate noise reduction techniques on the subsequent extraction of metrics from NDVI time series, a range of metrics was derived from the original, noise-reduced and noisy model time series. In this way a direct comparison of the resulting metrics was made between the six noise reduction strategies over different land cover types and at varying noise levels. An investigation of whether metric estimation from noise-reduced time series presented an improvement over estimation from unsmoothed time series was also included in the analysis. Table 3-4 lists the metrics selected for this evaluation.

Table 3-4: Summary of the NDVI time series metrics selected for evaluation.

Metric	Meaning/Description
Timing of SOS	Composite period in which growing season starts
Timing of EOS	Composite period in which growing season ends
LGS	Number of composite periods comprising the growing season
Maximum NDVI	Maximum level of photosynthesis reached during the growing season
Timing of Maximum NDVI	Composite period in which maximum level of photosynthesis is reached
NDVI Amplitude	Range between maximum level of photosynthesis, and average of pre- and post-growing season minimum levels of photosynthesis
Maximum Green-Up	Maximum rate of increase in photosynthetic activity during spring green-up
Timing of Maximum Green-Up	Composite periods between which maximum rate of increase in photosynthetic activity occurs
Average NDVI	Mean level of growing season photosynthesis
Integrated NDVI	Cumulative photosynthetic production during the growing season

A multitude of time series metrics is employed in the remote sensing of vegetation phenology, as evidenced by the review presented in Chapter Two. In order to capture the range of information that can be extracted from NDVI time series, a considerable variety of metrics was calculated for the present investigation. Not only do these include the most common metrics, such as start and length of the growing season, but also include other less common, yet still informative metrics, such as average growing season NDVI (Table 3-4). This provided a means of evaluating the effects of time series noise reduction on a variety of information extraction techniques. Descriptions of the selected metrics are given below. All metrics were derived for each

of the three years (2003 through 2005) contained in the six original model, 108 noise-reduced, and 18 noisy NDVI time series.

3.6.1.1 Start of Growing Season

As the most prevalent of NDVI time series metrics for the study of vegetation phenology, the start of growing season (SOS) metric was the first to be included in the present analysis (Table 3-4). It is important to estimations of growing season length, overall vegetative productivity and other phenological measures (DeFries *et* al. 1995, Chen *et al*. 1999, White *et al*. 1999). Because of its popularity, an investigation of the effects of noise reduction on the derivation of SOS from NDVI time series is important to the current evaluation of noise reduction techniques. Based on the review of SOS calculation methods presented in Chapter Two relative thresholding technique was selected for deriving SOS in the current analysis. It is more applicable and consistent between land cover types than an absolute threshold (Reed *et al*. 1994), and does not require the calculation of a function for each time series as do inflection point and curve derivative techniques (e.g. Zhang *et al*. 2003, Jönsson and Eklundh 2004). The latter techniques would further modify the original, noisy and noise-reduced time series during the calculation of SOS, and would result in less rigorous comparisons between noise reduction techniques.

A relative threshold of 10% of the NDVI range between the spring minimum and overall maximum NDVI was used to detect the point which the growing season had begun for each year (Figure 3-7). This was based on similar SOS calculations by Jönsson and Eklundh (2002, 2004), and was selected as an appropriate method for deriving SOS in the present case because of its flexibility across land cover types and ability to detect early spring vegetative growth at the surface. Spring minimum was identified as the minimum NDVI value between the beginning of the calendar year and the middle of the time series, assumed to represent the summer season. An NDVI of zero was used if this spring minimum was negative; negative values were assumed to be incorrect as they reflect an absence of vegetation (Myneni *et al*. 1995). The overall maximum NDVI was simply identified as the highest NDVI value in each yearly times series, and was assumed to reflect the highest level of surface vegetation photosynthesis reached during that

particular growing season (Figure 3-8). Because the date of each composite period represents the NDVI for the following 16 days, the composite period during which this SOS threshold was exceeded was identified as the timing of SOS.

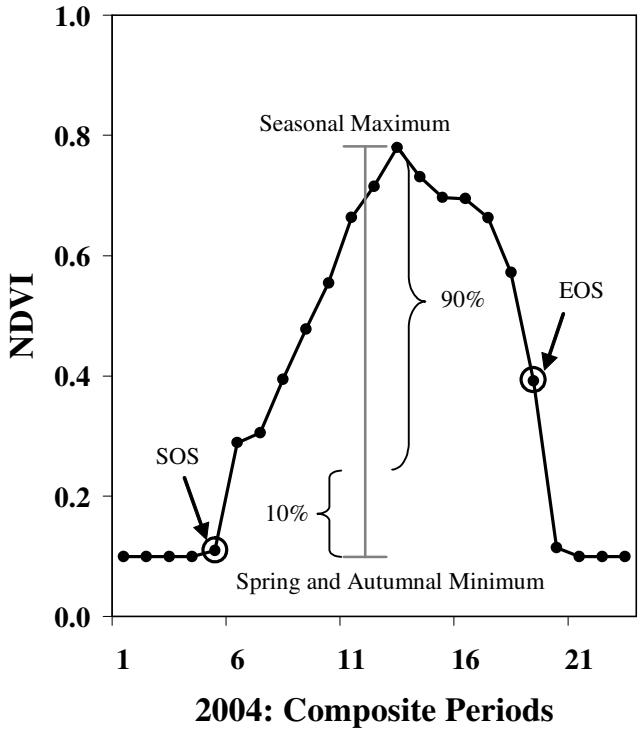

Figure 3-7: Illustration of SOS and EOS calculations, showing the thresholds used; in this case, these are identical because the spring and autumnal minimums are equal.

3.6.1.2 End of Growing Season

The end of growing season (EOS) metric, necessary for the determination of the length of the growing season, was derived for each year in a similar manner to SOS, again following the calculations of Jönsson and Eklundh (2002, 2004) (Table 3-4). A separate threshold of 90% of the NDVI range between overall maximum NDVI and the autumnal minimum NDVI was used to detect EOS for a particular year. As with spring minimum, the autumnal minimum was identified as the minimum NDVI value occurring in the latter half of a one-year time series, and was set to zero if this minimum was negative. The last composite period after the time of maximum NDVI, during which the NDVI value remains above this threshold, and after which it falls below, was determined

as the time of EOS (Figure 3-7). The selected composite period represents the last composite period of the growing season during which photosynthesis is occurring throughout.

3.6.1.3 Length of Growing Season

As a measure of the duration of photosynthesis and surface vegetation productivity, the length of the growing season (LGS) is not only an important phenological quality in itself, but forms the basis on which to estimate further metrics that characterize phenological processes occurring during the growing season (Table 3-4). LGS is often simply calculated as the time between SOS and EOS (Reed et al. 2003), and in this case, is recorded as the number of composite periods over which the growing season extends (Figure 3-9). That is, LGS includes both the SOS and EOS composite periods, as these represent periods over which NDVI exceeds the SOS or EOS thresholds, and with the number of composite periods occurring in between. This can be converted into days of the year, but because the minimum temporal resolution of the NDVI data set used in this is 16 days, this is not much more informative than recording the number of composite periods for the LGS metric.

3.6.1.4 Maximum NDVI

The level and timing of maximum NDVI in a one-year time series has been used to characterize the maximum levels of photosynthetic activity reached by surface vegetation over a variety of land cover types (e.g. Reed *et al*. 1994). It was simply calculated as the highest NDVI reached within a year-long time series (Table 3-4, Figure 3-8), and reflects the attainment of vegetation maturity at the surface (Jönsson and Eklundh 2002). The timing of maximum NDVI can be of importance in distinguishing between vegetated cover types at the surface (DeFries *et al*. 1995, Paruelo *et al*. 1998). Though simple to calculate, maximum NDVI and the timing of this maximum can be a powerful metric for characterizing vegetation growth and development.

3.6.1.5 NDVI Amplitude

NDVI amplitude provides a measure of vegetative productivity occurring over the course of a year. It relates to the amount of growth over one phenological cycle, and like maximum NDVI, can be important in the study of inter-annual phenological variation

and vegetative health, or land cover change (e.g. White *et al*. 2002). NDVI amplitude was estimated as the NDVI range between maximum NDVI and a minimum NDVI, calculated as the average of pre- and post- growing season minimum NDVI values and set to zero of this result was negative (Table 3-4, Figure 3-8). In this way, differences between winter NDVI minima before and after the growing season were accounted for. As was the case for all metrics, NDVI amplitude was calculated for each of the three years in each of the time series.

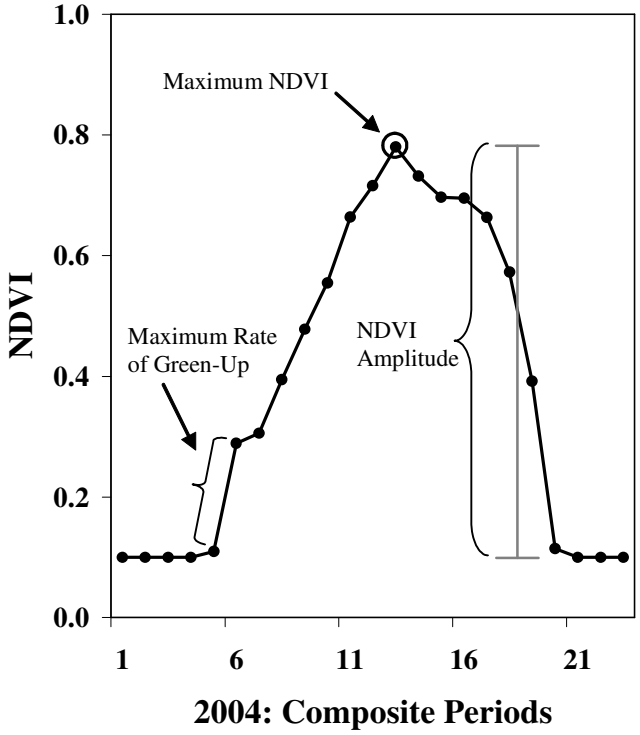

Figure 3-8: Illustration of maximum NDVI, maximum green-up and NDVI amplitude calculations.

3.6.1.6 Maximum Green-Up

As another measure of general phenological growth, the maximum rate of spring green-up and its timing reflect the period during which the highest rate of vegetative growth is occurring. While sometimes used as an indication of the start of the growing season (e.g. DeFries *et al*. 1995), this metric also provides a measure of symmetry or shape in a profile of phenological growth and development. By comparing the rate and timing of maximum green-up it is possible to detect whether the application of a noise

reduction algorithm to the time series has altered the shape of the profile during spring. The maximum rate of spring green-up was simply identified as the point at which the maximum increase in NDVI between two successive composite periods occurs (Table3-4, Figure 3-8). Both the rate of increase and the two composite periods between which it is found were recorded. Spring green-up rates are of greater importance in the literature than rates of senescence, as they are more reliably detected, modeled and predicted (Reed *et al* 1994, White *et al*. 2002, Wang *et al*. 2004, Kathuroju *et al*. 2007). Maximum senescence was not calculated in the current analysis.

3.6.1.7 Average Growing Season NDVI

Similar to maximum NDVI and NDVI amplitude, average growing season NDVI characterizes the levels of phenological growth occurring at the surface, but unlike these metrics is a measure of overall mean photosynthesis rather than the range in levels of photosynthesis. That is, while maximum NDVI and NDVI amplitude present an indication of the highest and lowest levels of detectable photosynthesis, average growing season NDVI indicates general levels. Simply calculated as the mean of all NDVI values occurring during the growing season (Moody and Johnson 2001), this metric provides additional information on vegetation health and development (Table 3-4, Figure 3-9).

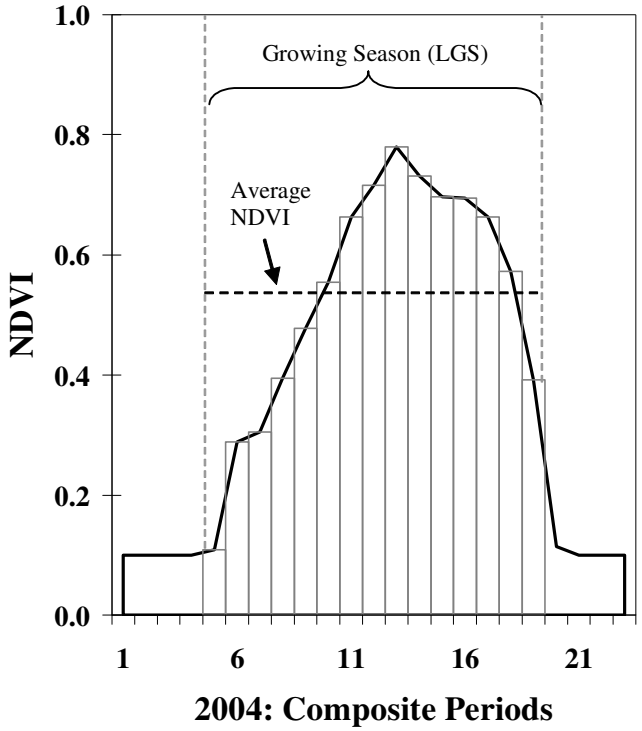

Figure 3-9: Illustration of average NDVI (dark dashed line) and I-NDVI calculations, showing the LGS (light grey dashed lines); the grey bars represent NDVI for the 16 days of each composite period in the growing season, which are summed to calculate I-NDVI.

3.6.1.8 Integrated NDVI

Growing season-integrated NDVI (I-NDVI) follows SOS, EOS and LGS in popularity for the remote sensing of vegetation phenology (Table 3-4). Shown to be closely related to measures of yearly vegetative net primary productivity (NPP) and seasonal biomass (Tucker *et al*. 1985, Goward *et al*. 1985, Prince 1991), this metric is an important phenological measure to a variety of ecological and climatological research objectives. It is typically calculated as the area under the NDVI curve between SOS and EOS (Justice *et al*. 1985, Reed *et al*. 1994, DeFries *et al*. 1995). Such a calculation, however, requires the presence of a function from which the integral can be calculated. As such a function does not exist for all NDVI times series in the current analysis except for those that have been smoothed by function-fitting, and because the fitting of a function through the data points in NDVI time series would further modify these curves, this was not done to calculate I-NDVI. Rather, an approximation of the area under the

NDVI curve for each growing season was made by calculating the sum of each NDVI value in the growing season multiplied by 16, to represent the 16 days included in the composite period represented by that value (Figure 3-9). In this way a general estimate of I-NDVI was made without altering the original, noisy and noise-reduced NDVI time series involved in the analysis.

3.6.2 Metric Data Processing

As was the case for the RMSE results described above, a number of descriptive statistics (i.e. mean, minimum, maximum, standard deviation) were calculated from the results of the metric calculations. For each metric, the calculations derived from each of the noise-reduced, original and noisy NDVI time series formed the basis of an overall, general statistical summary. Further analysis of these data is described below.

3.7 Data Analysis

It was recognized that in order to evaluate the overall performance of each noise reduction technique, a method for combining the results of the RMSE and metric calculations was needed. The following describes the calculations of raw and standardized performance scores, used not only to homogenize the two sets of evaluative data, but also to summarize and combine the results to produce informative comparisons between noise reduction techniques, and across land cover types and noise levels. The calculation of such scores also standardizes the various metric calculations, allowing for a comparison between the metrics themselves. Section 3.7.1 first describes the calculation of raw performance scores for the metric data set, as these methods formed the basis for the calculation of similar scores for the RMSE data set. Raw RMSE performance score calculations are presented in section 3.7.2, while section 3.7.3 explains the generation of standardized metric, RMSE and summary performance scores.

3.7.1 Metric Performance Score Calculation

Performance scores were devised to not only reflect the accuracy of the calculated metrics in relation to those derived from the original model NDVI time series, assumed to represent true, error-free metrics, but also to standardize these measures of accuracy across the various metrics. That is, the variation from one score to the next should

represent the variation from one metric calculation to another metric calculation, particularly in terms of variation in relation to the original model-derived metrics. In addition, because some metrics are calculated in terms of NDVI values and others are calculated in terms of a composite period (i.e. the timing of events), it is particularly important that performance scores be standardized across metrics. In light of these observations, the calculation of raw performance scores involved two steps. First, a normalization of the metric results was undertaken, followed by the actual raw score calculation itself.

Equation (3-10) was used to normalize the metric values to produce what could be referred to as a relative metric calculation.

$$RM_{lyn} = \frac{M_{lyn}}{MM_{ly}} \tag{3-10}$$

where RM_{lyn} is the relative metric for a particular land cover (l), year (y) and noise level (n), M_{lyn} is the metric calculated from a noisy or noise-reduced NDVI time series of the same land cover, year and noise level, and MM_{ly} is the original model time series-derived metric for the particular land cover and year. This calculation was performed for all six land covers (i.e. Grassland, Parkland, Boreal, Lower Foothills, Montane and Alpine), all three years in the time series (i.e. 2003, 2004 and 2005) and all three levels of introduced noise (10%, 40% and 70%). The result of Equation (3-10) reflects the proportion of the original model metric that each of the noise-reduced and noisy time series-derived metrics represents. For example, a standardized metric of 0.5 indicates a calculated metric is underestimating the true metric by 50%, while values greater than one indicate an overestimation. A standardized metric value of 1.00 indicates perfect agreement with the true model metric.

Once this normalization was performed, these standardized metric values were used to produce raw performance scores, calculated in such a way as to reflect increasing under- or overestimation with an increasing score. Table 3-5 shows the simple categorization technique employed in the performance score calculation.

Table 3-5: Categorization scheme used to calculate raw metric performance scores.

Relative Metric (RM)	Performance Score	Relative Metric (RM)	Performance Score
$0.0 \leq RM < 0.1$	10	$1.2 < RM \leq 1.3$	3
$0.1 \leq RM < 0.2$	9	$1.3 < RM \leq 1.4$	4
$0.2 \leq RM < 0.3$	8	$1.4 < RM \leq 1.5$	5
$0.3 \leq RM < 0.4$	7	$1.5 < RM \leq 1.6$	6
$0.4 \leq RM < 0.5$	6	$1.6 < RM \leq 1.7$	7
$0.5 \leq RM < 0.6$	5	$1.7 < RM \leq 1.8$	8
$0.6 \leq RM < 0.7$	4	$1.8 < RM \leq 1.9$	9
$0.7 \leq RM < 0.8$	3	$1.9 < RM \leq 2.0$	10
$0.8 \leq RM < 0.9$	2	$2.0 < RM \leq 2.1$	11
$0.9 \leq RM < 0.99$	1	⋮	⋮
$0.99 \leq RM \leq 1.01$	0	$2.9 < RM \leq 3.0$	20
$1.01 < RM \leq 1.1$	1	⋮	⋮
$1.1 < RM \leq 1.2$	2	$5.0 < RM \leq 6.0$	40

Raw performance score values range from zero, indicating metrics that are within one percent of the true model metric, to 40, indicating a 400% overestimation in metric calculation. While the production of such scores was undertaken in a relatively simple and straightforward manner, there is one assumption that should be made explicit. This is that variation in metric calculations has equal significance between metrics, particularly between NDVI value-based metrics and temporal, composite period-based metrics. For example, not only is it assumed that a 50% overestimation in maximum NDVI is as significant a variation in that metric as the same level of overestimation in I-NDVI, but that a similar overestimation in the timing of the start or end of season dates likewise represents as significant a deviation from the true start or end of season as does this deviation in other metrics. That is, in comparing the performance scores between metrics and across land covers, noise levels and years, one must assume that the meaning or weight of a particular score remains constant.

3.7.2 RMSE Performance Score Calculation

The calculation of raw metric performance scores provided the foundation for the similar calculation of raw RMSE performance scores. In order to allow for the comparison between standardized RMSE and metric performance scores, and the combination of these scores into the overall summary scores described in the following section, it was necessary to maintain a similar scale and range in raw scores between the two data sets. RMSE results are calculated in units of NDVI, and can therefore theoretically range between zero and one. It should be noted that while NDVI ranges between negative and positive one, RMSE cannot result in a negative number (see Equation 3-9). A comparison of metric performance scores and RMSE values showed that while the latter varied between approximately 0.02 and 0.3, all but one of the former ranged between 0.0 and 30.0.

Based on this information a simple strategy for calculating raw RMSE performance scores was followed. First, because the RMSE data set is much less complex than the metric data set in that only one RMSE value is calculated for each land cover type and noise level, and not per year or metric, and that all RMSE values share the same 'unit' – NDVI, though it must be remembered that NDVI is actually unitless – no standardization procedure was necessary for performance score calculation. Rather, these scores were calculated directly from the RMSE values themselves. This involved, first multiplying the RMSE by 100 to scale the data to that of the raw metric performance scores, and then second, rounding the decimal places of the values to generate integer scores like those calculated for the metrics. In this way, raw RMSE performance scores were calculated to the same scale and range as the metric scores, allowing for comparisons between the two, and their eventual combination into final performance scores.

The calculation of RMSE and metric performance scores relied on the assumption that any particular score represented the same deviation from the original model time series in both the RMSE and metric data sets. That is, it was assumed that a performance score of 5 would represent the same level of deviation from the original model time series in either case. Such an assumption was necessary for the direct comparison of RMSE

and metric calculation results, and the calculation of standardized and summary performance scores which comprised both measures of performance.

3.7.3 Standardized and Summary Performance Score Calculations

Once raw performance scores had been calculated for the metric and RMSE data sets, a series of performance score summarizations were constructed for each. This simply involved summing the raw performance scores for the set of noise-reduced and noisy time series separately, stratified by land cover type and noise level, and in the case of the metric scores, also by year and by metric. Total summations for all time series for each noise reduction technique and for all noisy time series were also made for both sets of raw performance scores.

These summed raw performance scores then provided the basis for standardized performance scores. Equation (3-11) was used to calculate standardized performance scores from the raw performance score summations for the RMSE and metric data sets, stratified by land cover, noise level, year and metric. Overall standardized performance scores were also calculated for the six noise reduction techniques and for all noisy time series, from raw performance scores totals calculated for each technique.

$$score_{Si} = \frac{score_i}{\left(\frac{\sum_{j=1}^{n} score_j}{n} \right)} \tag{3-11}$$

where $score_{Si}$ is the standardized score for the *i*th noise reduction technique for a particular category (i.e. a land cover, noise level, year or metric), $score_i$ is the total RMSE or metric raw performance score summation for the *i*th noise reduction technique for a particular category, $score_j$ is the raw performance scores of the *n* noise reduction techniques for a particular category, and *n* is the number of noise reduction techniques. It should be noted that raw summations and standardized performance scores for the noisy NDVI time series was also included in these calculations. The resulting standardized performance scores were not only used individually to evaluate the abilities of the six candidate noise reduction techniques to minimize noise and maintain data integrity, but

were also combined into standardized summary performance scores to produce an overall measure of performance for each. This simply involved summing the overall RMSE and metric standardized performance scores to produce a summary score for each technique. It is in this calculation of summary scores that the advantage of standardized performance scores is most evident. Because the total number of results comprising the RMSE data are considerably fewer than the number involved in the metric data (i.e. 18 RMSE values for each technique, versus 540 metric values), simply summing raw performance scores for these two data sets does not produce an adequate measure of overall performance; it would diminish the importance of the RMSE portion of the analysis in the final summary scores. However, the combination of standardized performance scores allows for equal representation of the results of the two evaluations in the final standardized summary score calculations. That is, in calculating these final standardized summary scores, it is assumed that the RMSE and metric portions of the analysis are of equal importance to an overall evaluation of performance, and that by summing standardized scores from the RMSE and metric performance score analyses, an overall measure of the general ability of each technique to reduce noise while maintaining data integrity, can be obtained.

3.7.4 Scenario Analysis

The last of the analyses comprising the present empirical comparison of selected noise reduction techniques for NDVI time series involved an investigation of how often and under what circumstances the application of noise reduction provides a benefit in relation to the lack of noise reduction. A simple procedure was developed in which the raw performance scores for each noise reduction technique and for each scenario were compared with the score calculated for the equivalent noisy time series. In other words, the number of scenarios in which a noise-reduced time series produces a better raw performance score than the equivalent noisy time series were evaluated in relation to the number of scenarios in which the noisy time series produced the same or a better raw performance score than a noise-reduced time series.

A scenario refers to the derivation of an RMSE result from one particular land cover with a particular level of noise and application of a particular noise reduction technique, or in the case of time series metrics, a metric result calculated for one

particular land cover, noise level and year. For instance, a scenario could comprise the calculation of RMSE from the MVI-filtered Grassland time series with 10% noise, or the calculation of SOS for the year 2003 of the MVI-filtered Grassland time series with 10% noise. The RMSE and metric results for these scenarios were compared with the RMSE and SOS calculated for the equivalent noisy time series – the noisy Grassland time series with 10% noise. The numbers of RMSE scenarios totalled 108 (i.e. six land covers, three noise levels and six noise reduction techniques), while the number of metric scenarios amounted to 3240 (i.e. six land covers, three noise levels, three years per time series, six noise reduction techniques and ten metrics). Summations or tallies of the numbers of scenarios demonstrating the benefit of noise reduction were compared with the numbers of scenarios that do not demonstrate a benefit to noise reduction. Total tallies were calculated for the sets of RMSE and metric scenarios, including one overall tally that comprised both RMSE and metric scenarios, as well as RMSE and metric scenario tallies stratified by land cover and noise level, and in the case of the latter set of scenarios, year and metric. Thus, the number of instances in which the application of noise reduction provides benefit and the situations in which this does or does not occur, could be evaluated. The results of these analyses are presented in the following chapter, as are the results of all analyses comprising the empirical comparison performed in the current research.

Chapter Four: Results

The results of the analyses described above are presented in the following chapter. Standardized performance score evaluation results are provided in section 4.1, where overall summary and stratified performance scores, along with the RMSE and metric standardized scores, are described. This is followed in section 4.2, by the results of the scenario analyses, which investigated the benefit of applying noise reduction. This latter section presents overall results, the RMSE and metric components of these overall results and the stratification of these results by land cover, noise level, and in the case of metric results, by year and by metric. It should be noted than in both sections 4.1 and 4.2, the seventh candidate noise reduction technique referred to represents the noisy data to which no noise reduction has been applied, often denoted in tables and figures as 'None'. The final section, section 4.3, provides a chapter summary. Additional results are provided in the appendices: Appendix C presents the NDVI time series generated by the six candidate noise reduction techniques; Appendix D provides the results of RMSE and metric calculations; and Appendix E comprises unstandardized performance scores.

4.1 Performance Evaluation: Standardized Performance Scores

Before examining the following results, it is necessary to comment on the meaning and limitations of these as measures of noise reduction technique performance. It must be recognized that the standardized performance scores calculated in the present analyses do not represent absolute, ratio-level quantifications of the performance of each selected noise reduction technique, and should not be interpreted as such. Rather, these results are interval-level data that provide relative measures of performance with regard to the other strategies under evaluation. The superiority or inferiority of performance demonstrated by the six candidate techniques must be understood as a comparative assessment; the following represents a relative rather than an absolute evaluation.

4.1.1 Overall Standardized Performance Scores

Figure 4-1 summarizes the final standardized summary performance scores for each of the candidate noise reduction techniques. All noise reduction techniques produced better standardized performance scores (1.78 to 2.10) than the noisy, unfiltered

NDVI time series (2.48). Of the six candidate techniques, the two function-fitting approaches, Double Logistic and Asymmetric Gaussian, performed best with scores of 1.78 and 1.79, respectively.

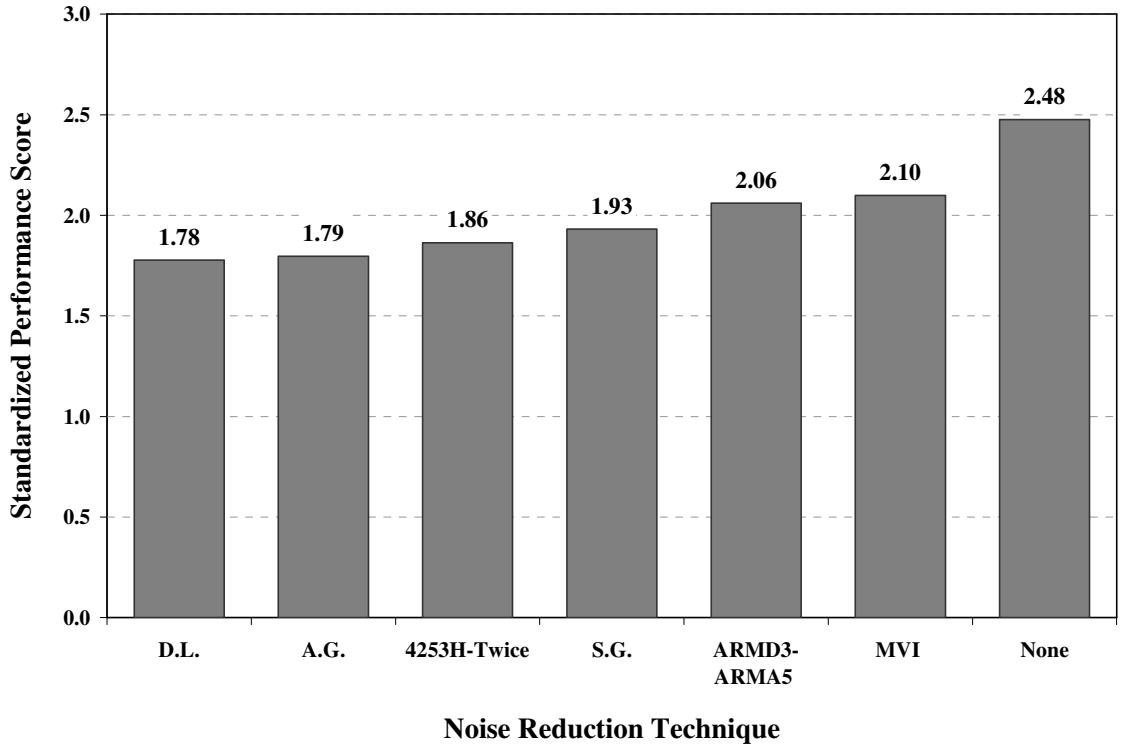

Figure 4-1: Standardized summary performance scores for the seven candidate noise reduction strategies, ordered from lowest (left) to highest (right). 'None' refers to the absence of noise reduction.

Figure 4-2 shows the overall root mean square error (RMSE) and metric standardized performance scores, illustrating their contributions to the total standardized summary scores. One can see that different patterns from those seen in Figure 4-1 emerge when the overall RMSE and metric performance scores are compared. For example, when evaluated strictly on the basis of overall RMSE scores alone, the Savitzky-Golay filter performed the best overall at minimizing noise, with a standardized RMSE score of 0.87. However, when measured by metric performance scores alone, the two function-fitting techniques (Double Logistic and Asymmetrical Gaussian) and 4253H-Twice filter were superior.

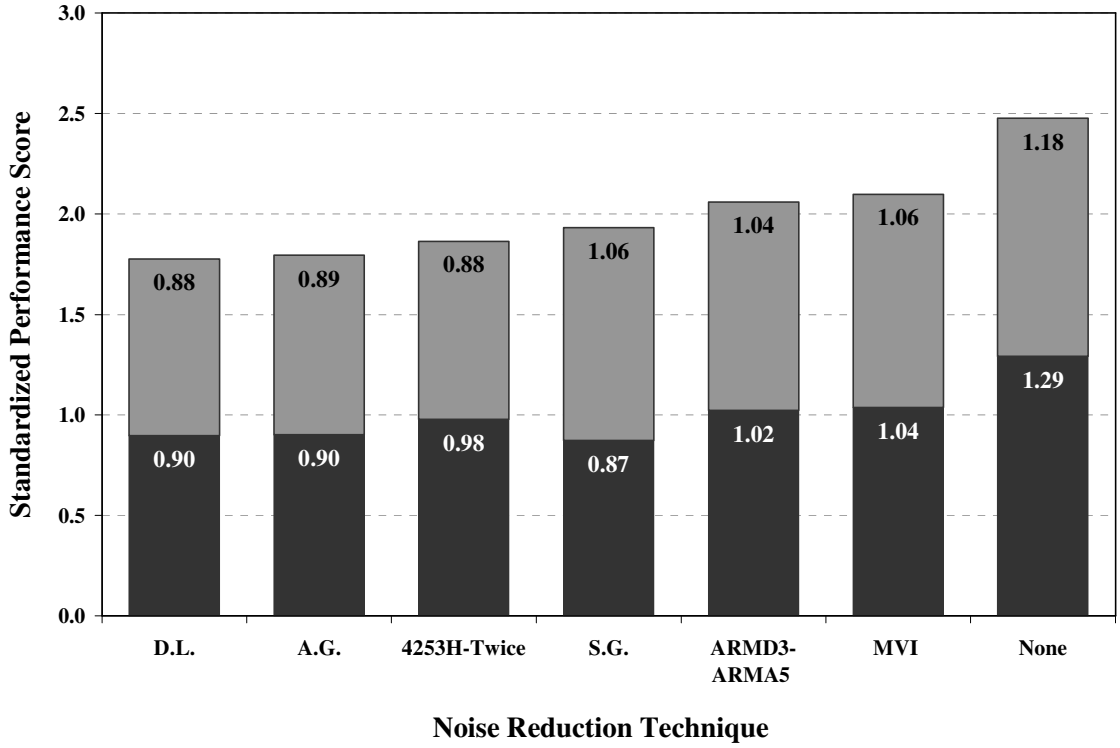

Figure 4-2: Overall standardized RMSE (dark grey) and metric (light grey) performance scores for the seven candidate noise reduction strategies, showing their respective contributions to the standardized summary performance scores. 'None' refers to the absence of noise reduction.

4.1.2 Stratified Performance Scores

While the overall performance score results provide an indicator of general metric efficiency across all the conditions modeled in this analysis, they do not provide insight into the specific influence of land cover type, noise level, or annual variation; nor do they reveal the impact of noise-reduction on individual phenological metrics. The results of analyses stratified by these factors are described below.

4.1.2.1 Performance Scores Stratified by Land Cover

Table 4-1 presents standardized summary performance scores stratified by land cover. More often than not, the raw noisy data generated the worst scores (up to 2.76), while of the candidate noise reduction techniques both the Double Logistic and Asymmetric Gaussian function-fitting techniques generally performed well with most scores below 2.0. The MVI and ARMD3-ARMDA5 filters performed poorly, often

generating scores above 2.0 (Table 4-1). Aside from these general patterns, however, considerable variation is observed in these results.

Table 4-1: Standardized summary performance scores for each of the seven candidate noise reduction techniques, stratified by land cover type. 'None' refers to the absence of noise reduction, and rankings per land cover are given in parenthesis.

Noise Reduction Technique	Grassland	Parkland	Boreal	Lower Foothills	Montane	Alpine
4253H-Twice Filter	1.87 (4)	1.73 (1)	1.74 (2)	2.04 (4)	1.52 (1)	2.08 (6)
ARMD3-ARMA5 Filter	2.27 (6)	2.09 (5)	2.09 (6)	2.06 (5)	1.74 (2)	2.01 (4)
Asymmetric Gaussian Function-Fitting	1.68 (2)	1.94 (3)	1.70 (1)	1.70 (2)	2.18 (5)	1.74 (2)
Double Logistic Function-Fitting	1.66 (1)	1.99 (4)	1.79 (3)	1.60 (1)	2.12 (4)	1.71 (1)
MVI Filter	2.26 (5)	2.15 (6)	2.00 (5)	2.23 (6)	1.81 (3)	2.06 (5)
Savitzky-Golay Filter	1.71 (3)	2.19 (7)	1.91 (4)	1.75 (3)	2.46 (7)	2.00 (3)
None	2.54 (7)	1.9 (2)	2.76 (7)	2.62 (7)	2.18 (6)	2.41 (7)

Tables 4-2 and 4-3 present the RMSE and metric standardized performance scores as stratified by land cover. One observes few similarities between these and the overall stratified results presented in Table 4-1; substantial variation exists within and across tables. While the technique producing the worst score for each land cover type remained almost constant across these three sets of results – the Savitzky-Golay filter for Parkland and Montane, and the raw, noisy data for Grassland, Boreal, Lower Foothills and Alpine – no additional trends or patterns arose.

Table 4-2: Standardized RMSE performance scores for each of the seven candidate noise reduction techniques, stratified by land cover type. 'None' refers to the absence of noise reduction, and rankings per land cover are given in parenthesis.

Noise Reduction Technique	Grassland	Parkland	Boreal	Lower Foothills	Montane	Alpine
4253H-Twice Filter	1.02 (3)	0.75 (1)	0.96 (4)	1.12 (5)	0.72 (1)	1.10 (4)
ARMD3- ARMA5 Filter	1.15 (4)	0.96 (2)	1.05 (6)	1.08 (4)	0.75 (2)	1.04 (3)
Asymmetric Gaussian Function-Fitting	0.80 (2)	1.07 (4)	0.84 (1)	0.82 (3)	1.18 (5)	0.86 (2)
Double Logistic Function-Fitting	0.80 (2)	1.07 (4)	0.90 (3)	0.76 (1)	1.21 (6)	0.83 (1)
MVI Filter	1.15 (5)	1.07 (4)	1.02 (5)	1.12 (6)	0.75 (3)	1.04 (3)
Savitzky-Golay Filter	0.78 (1)	1.07 (4)	0.87 (2)	0.80 (2)	1.24 (7)	0.86 (2)
None	1.29 (6)	1.02 (3)	1.35 (7)	1.32 (7)	1.14 (4)	1.28 (5)

Table 4-3: Standardized metric performance scores for each of the seven candidate noise reduction techniques, stratified by land cover type. 'None' refers to the absence of noise reduction, and rankings per land cover are given in parenthesis.

Noise Reduction Technique	Grassland	Parkland	Boreal	Lower Foothills	Montane	Alpine
4253H-Twice Filter	0.84 (1)	0.98 (4)	0.78 (1)	0.93 (3)	0.80 (1)	0.98 (3)
ARMD3- ARMA5 Filter	1.13 (6)	1.12 (6)	1.04 (6)	0.99 (5)	0.98 (3)	0.97 (2)
Asymmetric Gaussian Function-Fitting	0.88 (3)	0.87 (1)	0.86 (2)	0.88 (2)	1.00 (4)	0.88 (1)
Double Logistic Function-Fitting	0.85 (2)	0.92 (3)	0.89 (3)	0.84 (1)	0.91 (2)	0.88 (1)
MVI Filter	1.12 (5)	1.08 (5)	0.98 (4)	1.11 (6)	1.06 (6)	1.02 (4)
Savitzky-Golay Filter	0.93 (4)	1.12 (7)	1.04 (5)	0.95 (4)	1.22 (7)	1.14 (6)
None	1.25 (7)	0.89 (2)	1.41 (7)	1.30 (7)	1.04 (5)	1.13 (5)

4.1.2.2 Performance Scores Stratified by Noise Level

Tables 4-4, 4-5 and 4-6 list standardized summary, RMSE and metric performance scores for the seven candidate noise reduction strategies, stratified by noise level, respectively. Three principal observations should be made. First, in all three sets of results, no one technique out-performed all other techniques across the 10%, 40% and 70% noise levels; the performance of each varied by noise level. Second, at moderate to high noise levels (40% and 70%) the raw, unfiltered NDVI time series generated the worst standardized summary (2.95 and 2.63), RMSE (1.37 and 1.28) and metric (1.58 and 1.36) performance scores, but at slight noise levels (10%) the Savitzky-Golay filter performed the poorest according to all three sets of scores (2.54, 1.24 and 1.30 as summary, RMSE and metric scores) (Tables 4-4, 4-5 and 4-6). Thirdly, as exemplified by the numbers just given, the raw, noisy data generated worse standardized performance scores at the 40% noise level than at the 70% noise level. The latter is unexpected, as one would undoubtedly anticipate the worst scores to be produced at the 70% noise level.

Table 4-4: Standardized summary performance scores for each of the seven candidate noise reduction techniques, stratified by noise level. 'None' refers to the absence of noise reduction, and rankings per noise level are given in parenthesis.

Noise Reduction Technique	10% Noise	40% Noise	70% Noise
4253H-Twice Filter	1.68 (1)	1.76 (3)	1.99 (3)
ARMD3-ARMA5 Filter	2.03 (5)	1.94 (5)	2.11 (4)
Asymmetric Gaussian Function-Fitting	1.94 (4)	1.76 (2)	1.65 (1)
Double Logistic Function-Fitting	1.89 (2)	1.70 (1)	1.70 (6)
MVI Filter	1.97 (3)	2.04 (6)	2.16 (5)
Savitzky-Golay Filter	2.54 (6)	1.86 (4)	1.75 (2)
None	1.94 (4)	2.95 (7)	2.63 (7)

Table 4-5: Standardized RMSE performance scores for each of the seven candidate noise reduction techniques, stratified by noise level. 'None' refers to the absence of noise reduction, and rankings per noise level are given in parenthesis.

Noise Reduction Technique	10% Noise	40% Noise	70% Noise
4253H-Twice Filter	0.85 (1)	0.94 (4)	1.06 (4)
ARMD3-ARMA5 Filter	1.00 (4)	0.97 (5)	1.07 (5)
Asymmetric Gaussian Function-Fitting	0.98 (3)	0.91 (3)	0.86 (2)
Double Logistic Function-Fitting	0.98 (3)	0.89 (2)	0.87 (3)
MVI Filter	0.93 (2)	1.06 (6)	1.07 (5)
Savitzky-Golay Filter	1.24 (6)	0.86 (1)	0.80 (1)
None	1.02 (5)	1.37 (7)	1.28 (6)

Table 4-6: Standardized metric performance scores for each of the seven candidate noise reduction techniques, stratified by noise level. 'None' refers to the absence of noise reduction, and rankings per noise level are given in parenthesis.

Noise Reduction Technique	10% Noise	40% Noise	70% Noise
4253H-Twice Filter	0.83 (1)	0.82 (2)	0.93 (3)
ARMD3-ARMA5 Filter	1.03 (5)	0.97 (4)	1.05 (5)
Asymmetric Gaussian Function-Fitting	0.97 (4)	0.84 (3)	0.79 (1)
Double Logistic Function-Fitting	0.91 (2)	0.81 (1)	0.83 (2)
MVI Filter	1.04 (6)	0.98 (5)	1.09 (6)
Savitzky-Golay Filter	1.30 (7)	1.00 (6)	0.95 (4)
None	0.92 (3)	1.58 (7)	1.36 (7)

4.1.2.3 Performance Scores Stratified by Year

This section pertains only to standardized metric performance score results and does not include standardized summary or RMSE performance scores. This is because the RMSE calculations comprised a single three-year-long time series and were not calculated on a per-year basis as were the metric derivations, and therefore could not be stratified by year. Without a set of metric and RMSE results to combine, summary results were also not possible. Nevertheless, a comparison of standardized metric

performance scores, stratified by year, still provided useful insight into the factors influencing noise reduction technique performance.

Table 4-7 presents the standardized metric performance scores stratified by year. One can observe variation in these scores between the three years, but there are two general trends to note. First, the numbers show that the Double Logistic and Asymmetric Gaussian function-fitting approaches and the 4253H-Twice filter all produced consistently better performance scores (≤ 0.91) than the remaining three filters and the unfiltered NDVI time series, which thus consistently performed more poorly and produced worse scores (≥ 0.98) (Table 4-7). Second, further examination of Table 4-7 reveals that a greater range exists between standardized metric performance scores for any particular year than between scores for any particular noise reduction technique across the three years. That is, more variation exists between the performances of the candidate techniques than exists for one technique across time.

Table 4-7: Standardized metric performance scores for each of the seven candidate noise reduction techniques, stratified by year. 'None' refers to the absence of noise reduction, and rankings per year are given in parenthesis.

Noise Reduction Technique	2003	2004	2005
4253H-Twice Filter	0.91 (2)	0.84 (1)	0.90 (1)
ARMD3-ARMA5 Filter	1.10 (5)	0.98 (4)	1.03 (4)
Asymmetric Gaussian Function-Fitting	0.91 (2)	0.88 (3)	0.90 (1)
Double Logistic Function-Fitting	0.85 (1)	0.87 (2)	0.91 (2)
MVI Filter	1.06 (3)	1.12 (6)	1.01 (3)
Savitzky-Golay Filter	1.11 (6)	1.06 (5)	1.01 (3)
None	1.06 (4)	1.24 (7)	1.24 (5)

4.1.2.4 Performance Scores Stratified by Metric

As with the previous set of results, only standardized metric performances scores could be stratified by metric, and the following section therefore includes neither RMSE nor summary score results; it is nevertheless informative. Table 4-8 lists the standardized metric performance scores stratified by metric, which appear to contain greater variation than is observed with stratification by land cover, noise level or year. For example, the

Savitzky-Golay filter produces the best (0.53) and worst (1.75) standardized performance scores in Table 4-8. Two general observations may be made, however. First, one can observe that for all but the timing of maximum NDVI, the Double Logistic and Asymmetric Gaussian function-fitting techniques out-performed both the MVI and ARMD3-ARMA5 filters. The performances of the 4253H-Twice and Savitzky-Golay filters, and the raw, noisy data, were much more variable with metric. Second, one can observe that for four of the five time-based metrics – SOS, EOS, LGS and the timing of maximum green-up – refraining from the application of noise reduction produced the worst standardized performance scores. For the remaining metrics, at least two of the six candidate noise reduction techniques performed more poorly than the unfiltered, noisy data (Table 4-8).

Table 4-8: Standardized metric performance scores for each of the seven candidate noise reduction techniques, stratified by metric. 'None' refers to the absence of noise reduction, and rankings per metric are given in parenthesis.

Metric	4253H-Twice	ARMD3-ARMA5	A.G.	D.L.	MVI	S.G.	None
Maximum NDVI	0.97(4)	1.39(5)	0.92(3)	0.92(3)	1.39(5)	0.53(1)	0.86(2)
Time of Maximum NDVI	0.75(1)	0.90(4)	1.23(6)	1.35(7)	0.79(2)	0.85(3)	1.13(5)
NDVI Amplitude	0.94(3)	1.16(6)	0.81(2)	0.78(1)	1.08(5)	1.24(7)	0.99(4)
Start of Season	0.86(1)	0.97(4)	0.95(3)	0.92(2)	0.99(5)	1.14(6)	1.17(7)
End of Season	0.75(2)	1.15(5)	0.86(3)	0.90(4)	1.21(6)	0.68(1)	1.46(7)
Length of Season	0.83(1)	0.98(4)	0.84(3)	0.83(2)	1.10(5)	1.02(5)	1.41(7)
Maximum Green-Up	0.79(1)	0.94(5)	0.89(4)	0.85(3)	0.96(6)	1.75(7)	0.81(2)
Time of Maximum Green-Up	0.99(4)	0.98(3)	0.90(2)	0.89(1)	1.02(5)	1.02(5)	1.21(6)
Integrated NDVI	0.95(2)	1.07(5)	0.92(1)	0.92(1)	1.13(6)	1.03(4)	0.96(3)
Average NDVI	0.97(4)	1.22(5)	0.71(2)	0.64(1)	1.26(6)	1.35(7)	0.85(3)

4.2 Scenario Analysis Results

The scenario analysis described in the previous chapter was designed to investigate when and in what circumstances the application of noise reduction provided a

benefit over the raw, noisy NDVI time series. The following presents the overall and stratified results of the scenario tallies calculated in the analysis, which include summations of both RMSE and metric scenarios. Several of these show potentially contradictory results to what was observed regarding standardized performance scores, particularly the overall results and those results pertaining to metric scenarios.

The overall scenario tallies, comprising all RMSE and metric scenarios, are presented in Figure 4-3. Of a total 3348 scenarios, 2228 or 67% showed better or similar performance scores when noise reduction had not been applied, while the remaining 1120 or 33% showed improvement in performance scores with the application of noise reduction. Thus, in the majority of all situations comprising the current analysis, applying noise reduction did not produce a benefit.

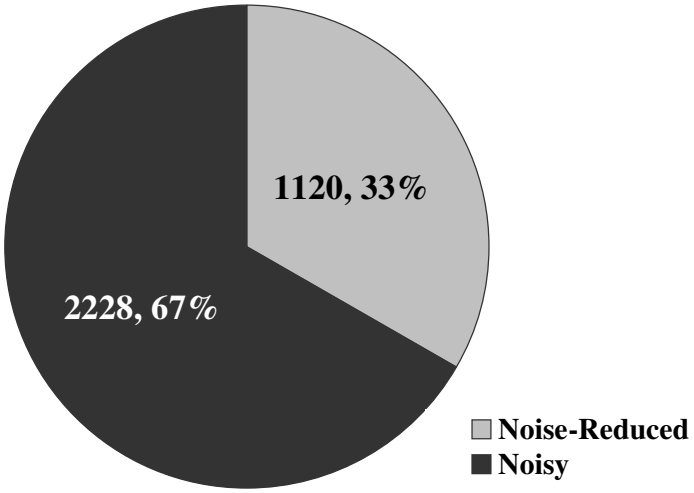

Figure 4-3: Pie chart showing the number and percentage of scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score.

Figure 4-4 presents separate metric and RMSE scenario tally results, which comprised the overall results presented in the previous figure. The majority of metric scenarios showed no benefit from noise reduction, with 68% of scenarios demonstrating the same or better performance scores from noisy, unfiltered time series as from noise-reduced time series. However, the majority (73%) of RMSE scenarios generated better performance scores when noise reduction had been applied.

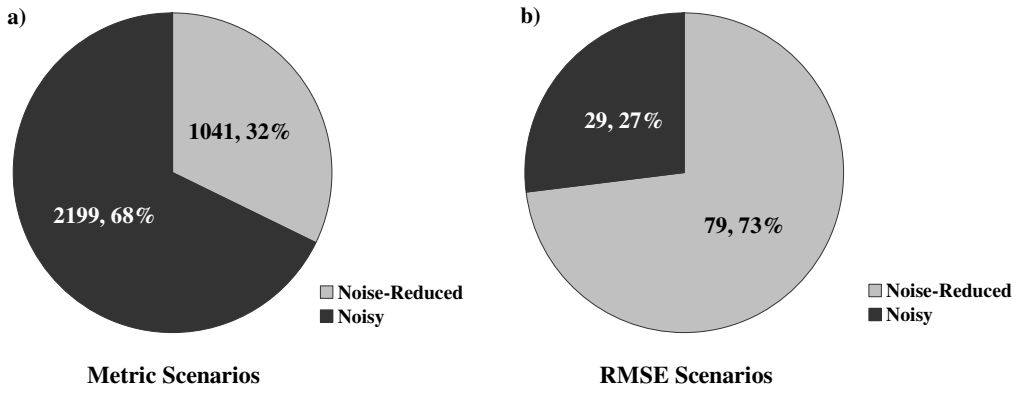

Figure 4-4: Pie chart showing the number and percentage of scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score for a) metric and b) RMSE scenario results.

Stratification of the RMSE scenario results by noise reduction technique are presented in Figure 4-5. Similar to the overall RMSE scenario results, the application of noise reduction provided benefit in most cases (67% to 83%) for every noise reduction technique. Indeed, the RMSE results do not vary much between techniques.

Figure 4-6 presents metric scenario results, stratified by noise reduction technique. As is observed for the RMSE scenarios, the numbers do not vary much between noise reduction techniques; for each, between 32% and 44% of metric scenarios demonstrated better performance scores with the application of noise reduction. These reflect the overall metric scenario findings presented in Figure 4-4.

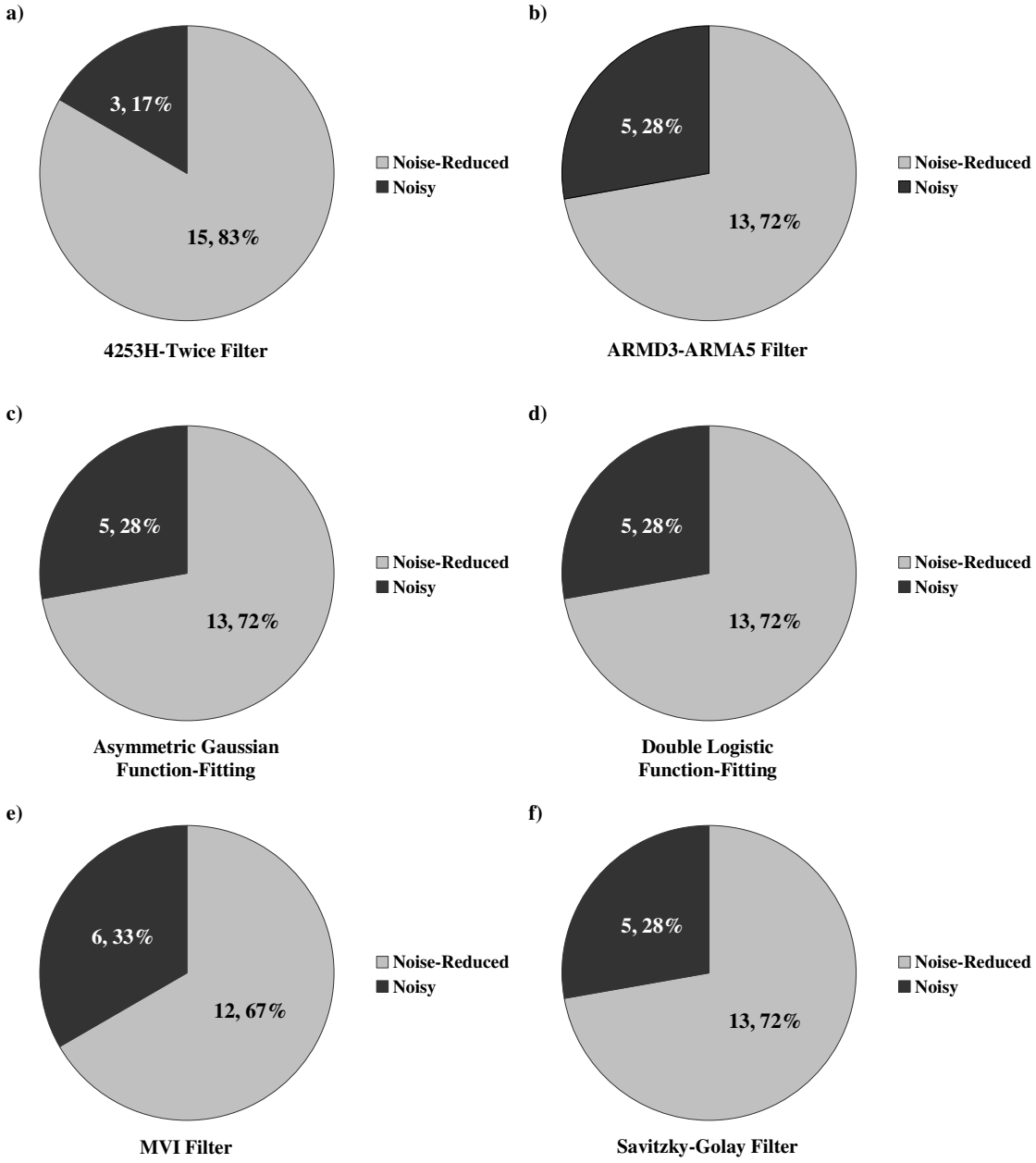

Figure 4-5: Numbers and percentages of RMSE scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the a) 4253H-Twice filtering, b) ARMD3-ARMA5 filtering, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filtering and f) Savitzky-Golay filtering techniques.

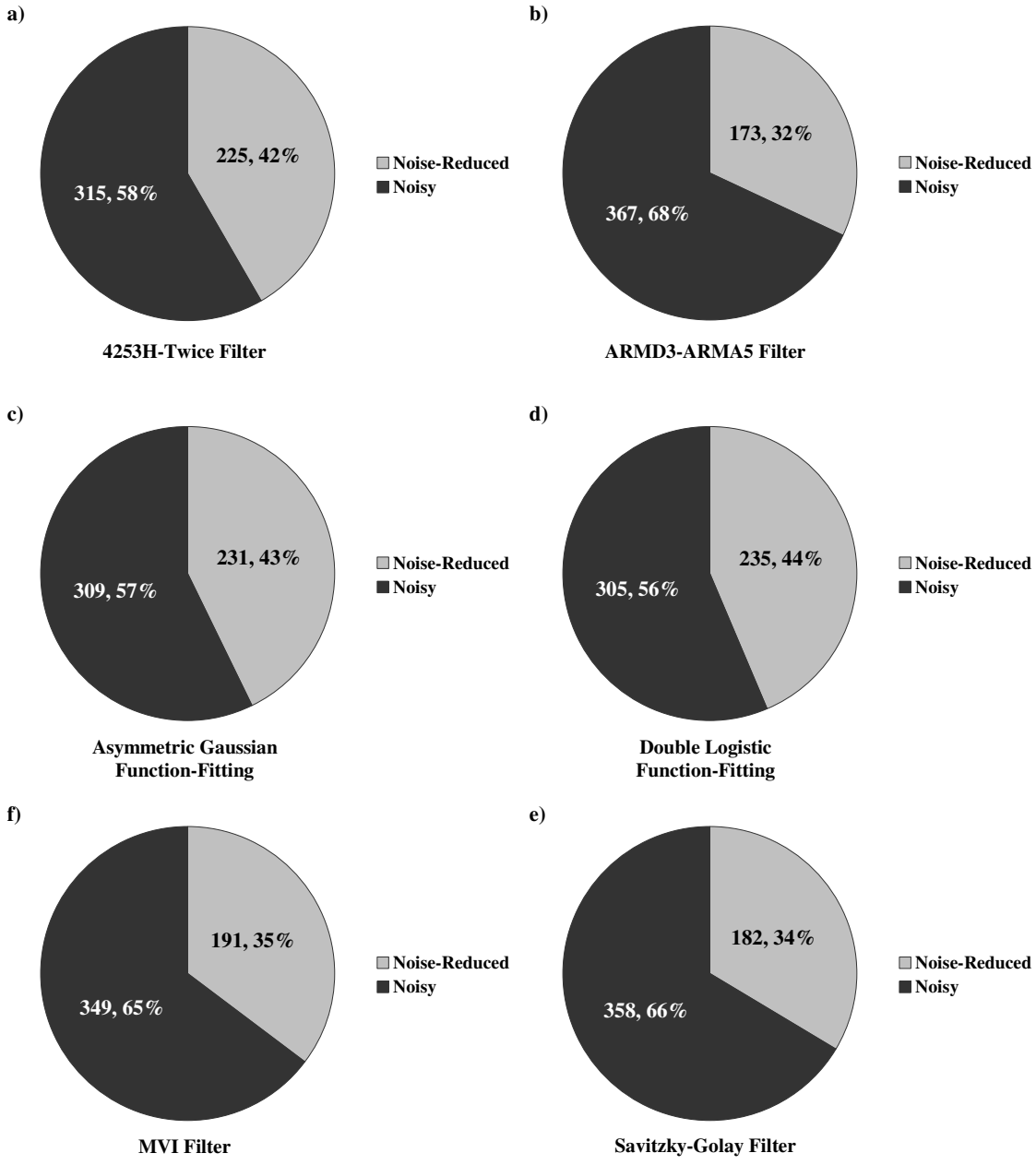

Figure 4-6: Numbers and percentages of metric scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the a) 4253H-Twice filtering, b) ARMD3-ARMA5 filtering, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filtering and f) Savitzky-Golay filtering techniques.

Figure 4-7 shows the stratification of RMSE scenario tallies by land cover. The most striking aspect of these results is their considerable variability, particularly when compared to the stratification of scenarios by noise reduction technique. The percentages

of scenarios which generated better results with the application of noise reduction range from 50% for the Parkland and Montane land covers, to 100% for the Boreal land cover.

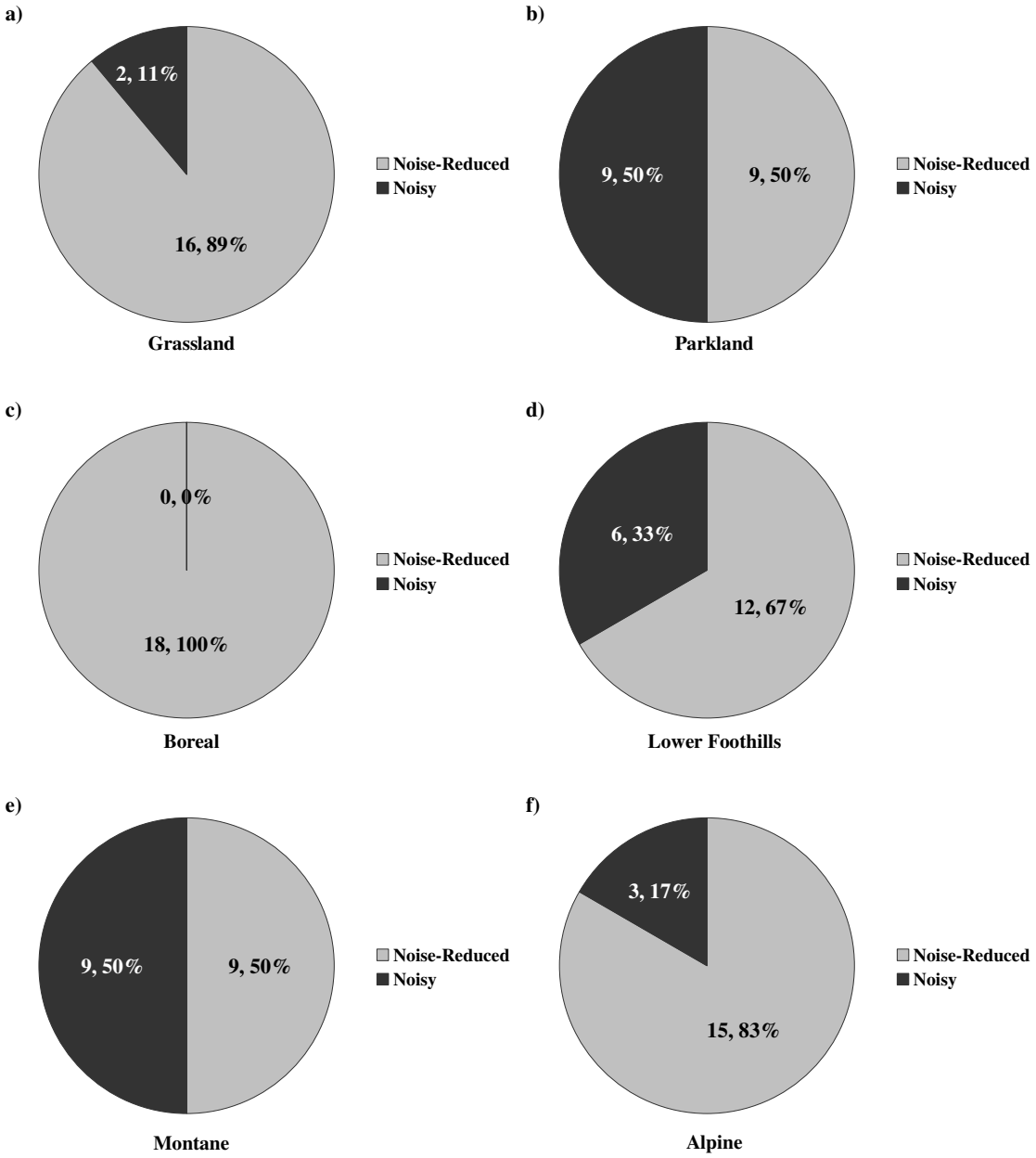

Figure 4-7: Numbers and percentages of RMSE scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the a) Grassland, b) Parkland, c) Boreal, d) Lower Foothills, e) Montane and f) Alpine land covers.

The metric scenario results, stratified by land cover, show more similarity across land covers than the RMSE scenario results, as observed in Figure 4-8. Comparable to

the overall metric tallies, the percentages of metric scenarios having generated better performance scores with noise reduction range from 31% for the Parkland land cover to 46% for the Lower Foothills land cover.

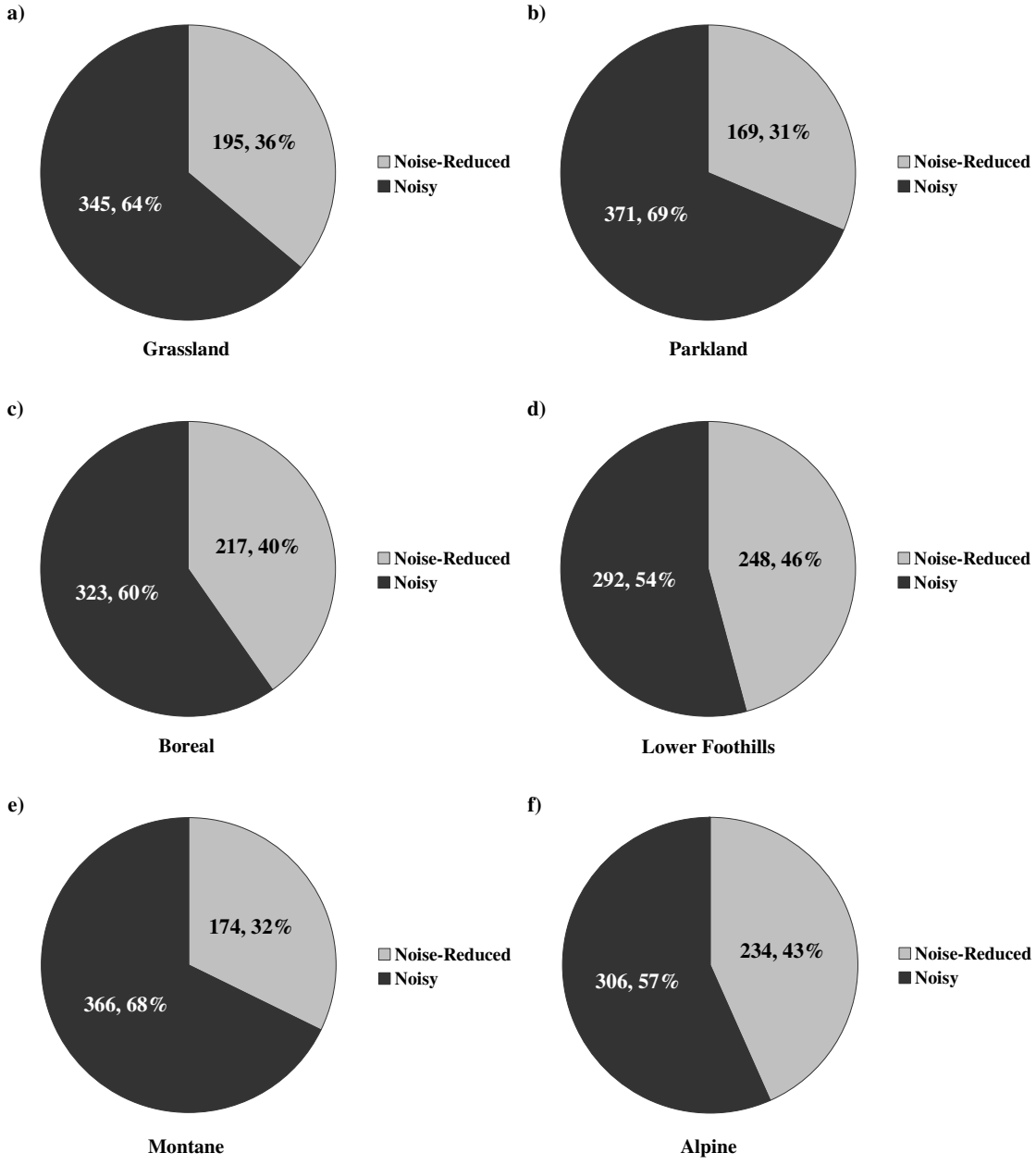

Figure 4-8: Numbers and percentages of metric scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the a) Grassland, b) Parkland, c) Boreal, d) Lower Foothills, e) Montane and f) Alpine land covers.

Figure 4-9 presents the RMSE scenario tallies, stratified by noise level. At both the 40% and 70% noise levels the great majority of scenarios (92% and 86%, respectively) demonstrated improved performance scores with the application of a noise reduction technique. However, at the 10% noise level, this dropped to just 42%.

A similar pattern to that seen in the RMSE scenario results is observed for the metric scenario results when stratified by noise level, though different percentages were obtained (Figure 4-10). The 40% and 70% noise levels show very similar numbers, with 45% and 44% of scenarios having demonstrated a benefit with noise reduction, respectively, while at the 10% noise level this percentage drops to 25%.

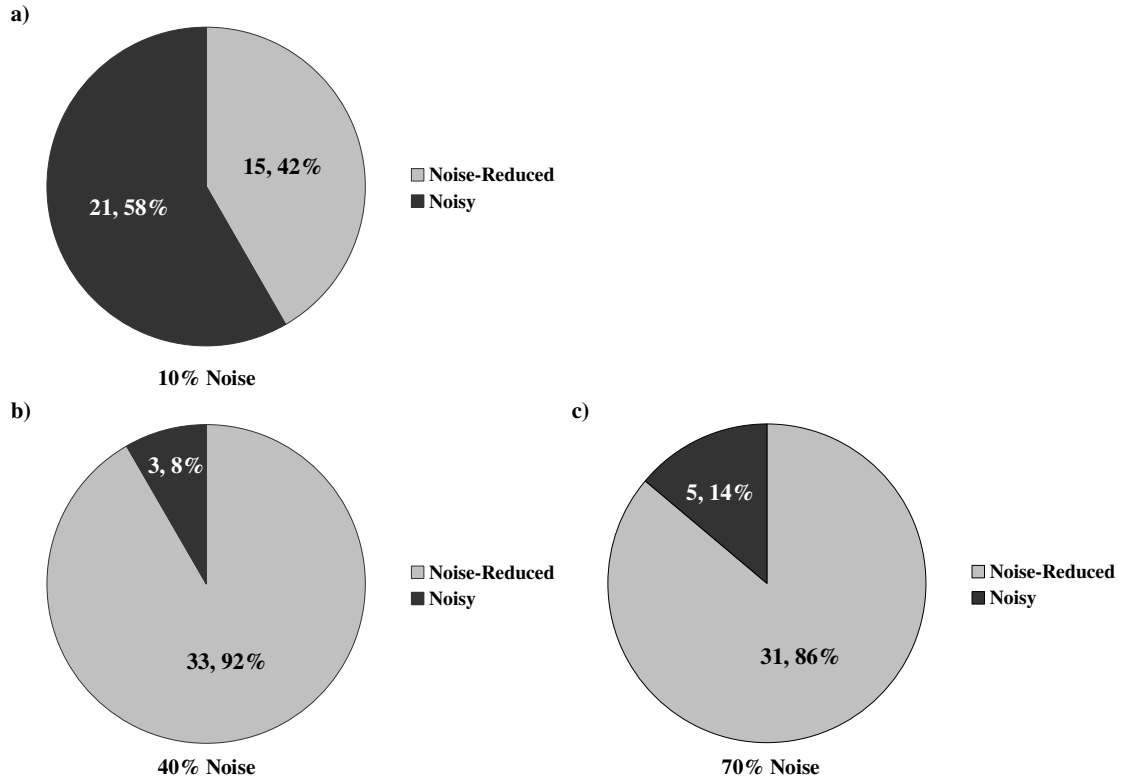

Figure 4-9: Numbers and percentages of RMSE scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by noise at the levels of a) 10%, b) 40% and c) 70%.

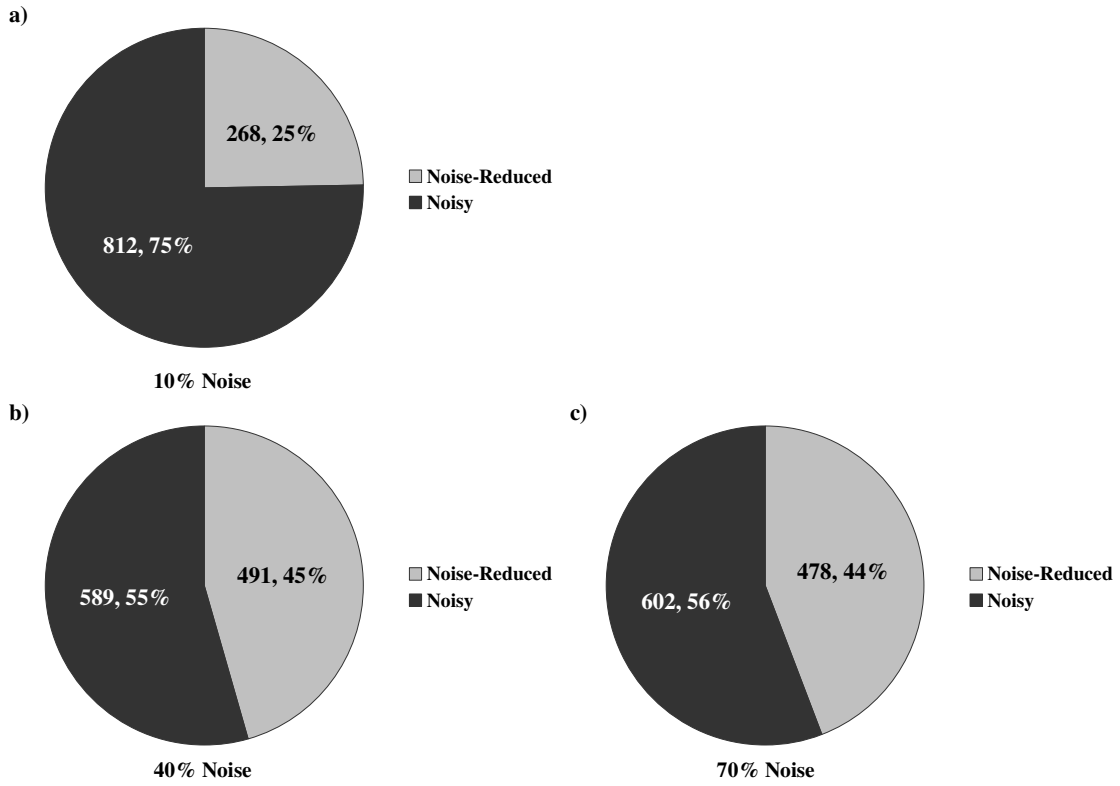

Figure 4-10: Numbers and percentages of metric scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by noise at the levels of a) 10%, b) 40% and c) 70%.

The final two figures again pertain only to metric scenario results, as no equivalent RMSE result stratifications are possible on the basis of year or metric. These figures are nonetheless very informative. Figure 4-11 presents the metric scenario tallies stratified by year. It is immediately apparent that the results vary little between 2003, 2004 and 2005, showing that 36%, 37% and 42% of the metric scenarios, respectively, produced improved performance scores with the application of a noise reduction technique. These, again, reflect the overall metric scenario tallies presented in Figure 4-3. Figure 4-12, which shows the metric scenario tallies stratified by metric, demonstrates considerably more variability, however. The number of metric scenarios that demonstrated the benefits of noise reduction ranged from just 15% for maximum NDVI, to 61% for the maximum rate of green-up. In addition to this variability, another, perhaps more interesting observation must be made. That is, that three of the metrics, NDVI amplitude, maximum rate of green-up and average NDVI all showed that in more than

half of these metric scenarios (56%, 61% and 59%, respectively), applying noise reduction was beneficial. These are the only metric scenario results to suggest that noise reduction is more often beneficial than relying on the raw, noisy NDVI time series for metric derivation.

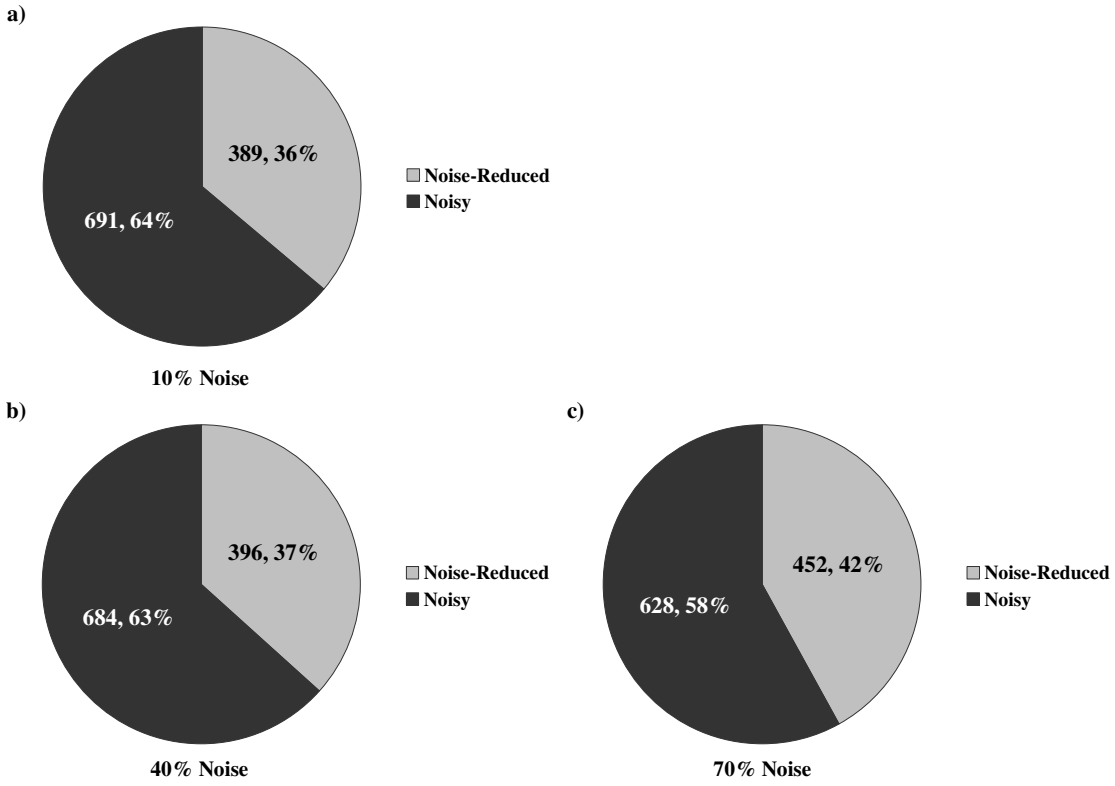

Figure 4-11: Numbers and percentages of metric scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the years a) 2003, b) 2004 and c) 2005.

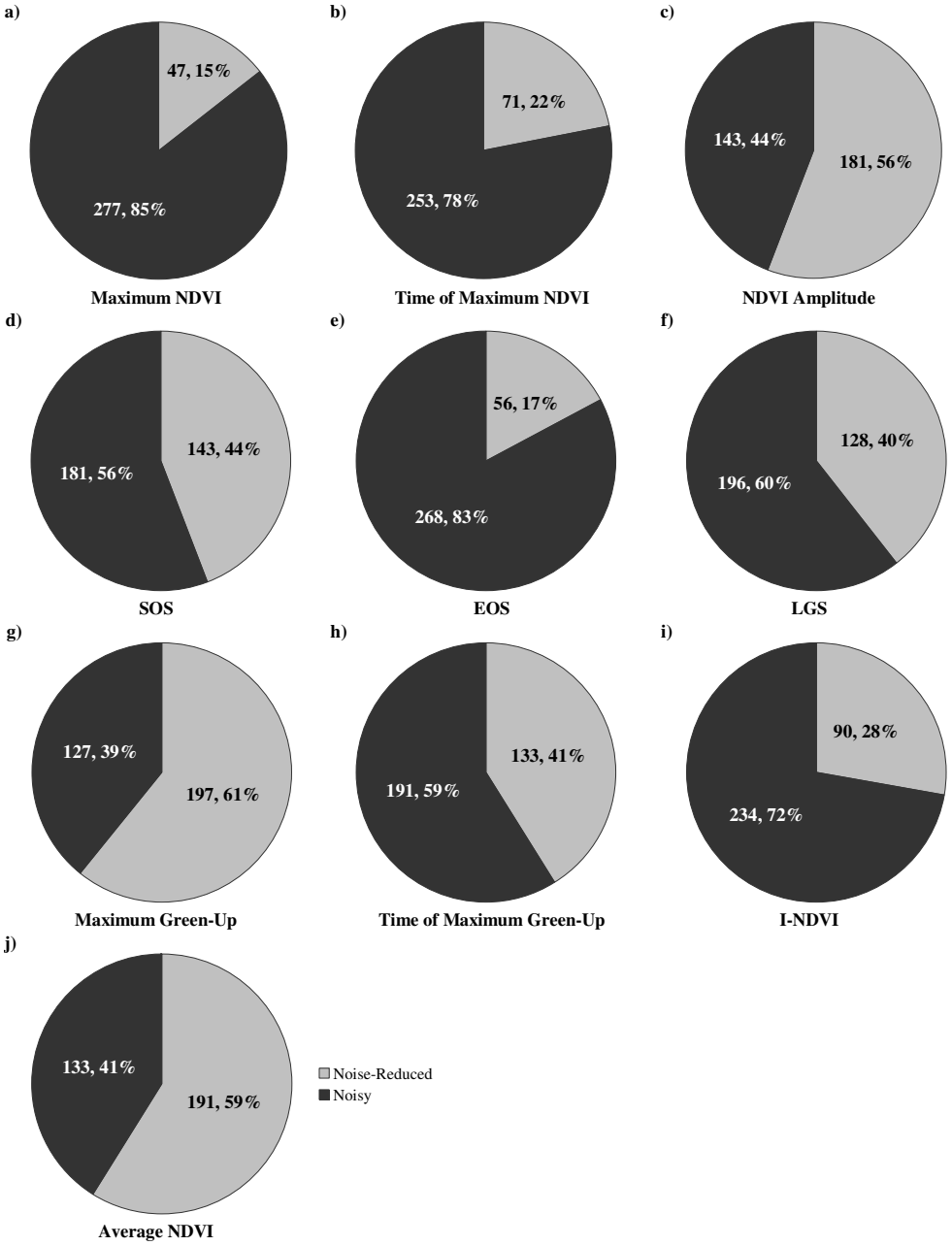

Figure 4-12: Numbers and percentages of metric scenarios in which either a noise-reduced time series (light grey) or the noisy, unfiltered times series (dark grey) produced a better unstandardized performance score, stratified by the metrics a) maximum NDVI, b) timing of maximum NDVI, c) NDVI amplitude, d) SOS, e) EOS, f) LGS, g) maximum green-up, h) timing of maximum green-up, i) I-NDVI, and j) average NDVI.

4.3 Chapter Summary

Overall standardized summary performance score results indicated that the two function-fitting techniques, the Double Logistic and Asymmetric Gaussian, performed the best of the six candidate noise reduction techniques, while the raw, unfiltered data produced the worst overall performance. The RMSE and metric components of these summary scores differed, however; the Savitzky-Golay filter generated the best overall RMSE results, with the two function-fitting methods and the 4253H-Twice filter produced the best overall metric results.

Stratification of the results revealed that standardized summary, RMSE and metric performance scores varied considerably with land cover. The consistently poor performance by the unfiltered, noisy data across the majority of land covers was the one observable trend. Similarly, no one technique performed the best across all three levels of introduced noise according to the stratified results. Nevertheless the standardized summary, RMSE and metric performance scores showed that not applying noise reduction produced the worst performance scores at the 40% and 70% noise level, while at the 10% noise level the Savitzky-Golay filter performed poorest. An unexpected result was also encountered in these stratified performance scores. It was observed that the unfiltered data with 40% noise generated worse standardized performance scores than the unfiltered data with 70% noise. The opposite would have been expected.

Stratification by year and metric was necessarily restricted to standardized metric performance scores; RMSE and summary results were not included in this portion of the analysis. According to metric performance scores stratified by year, more variation in scores occurred between the candidate noise reduction techniques for each year, than between the three years for each technique. In addition, the Double Logistic and Asymmetric Gaussian function-fitting methods, and the 4253H-Twice filter consistently produced better standardized metric performance scores than the MVI, ARMD3-ARMA5 and Savitzky-Golay filters across the years 2003 through 2005. The stratification by metric of standardized metric performance scores revealed more variation in results between metrics than was observed for land cover, noise level or year. Indeed, the general superiority of the Double Logistic and Asymmetric Gaussian function-fitting

methods over the MVI and ARMD3-ARMA5 filters across the majority of metrics was the only observable pattern in these results.

Overall scenario tallies from the scenario analysis showed that in the majority of all cases, applying a noise reduction technique did not improve on the raw performance scores produced from raw, noisy data. This was also demonstrated by the overall metric scenario tallies, but the overall RMSE results showed a contrary pattern. This dichotomy between RMSE and metric results was generally consistent. In the majority of RMSE scenarios, applying noise reduction produced an improvement in performance scores. RMSE and metric scenario tallies varied little when stratified by noise reduction technique, but RMSE scenario results varied considerably with land cover. Metric scenario results showed more variation with land cover than with noise reduction technique, but not to the same degree as that demonstrated by the RMSE scenarios. Stratification by noise level revealed that noise reduction is consistently less beneficial at the 10% noise level than at the 40% and 70% noise levels, according to both RMSE and metric scenario tallies. Stratification of metric scenario results by year produced little variation in the results; tallies with the majority of scenarios demonstrating no improvement in performance scores with noise reduction were consistently calculated for all three years. The final stratification, that of the metric scenario results by metric, revealed considerable variation. No discernable trends were found in these results, although three of the metrics generated tallies contrary to all other metric scenario results. More than half of the NDVI amplitude, maximum rate of green-up and average NDVI scenarios demonstrated improved performance scores with the application of noise reduction.

The results presented in the current chapter enable the informal research questions posed in Chapter One to be answered; they support the successful completion of the four objectives outlined in that chapter. These questions are explicitly addressed in the following chapter through a discussion of the above results and their implications for the reduction of noise in NDVI time series.

Chapter Five: Discussion

The following chapter is organized in a manner similar to that found in the previous chapter; it is structured around the main objectives outlined in the introductory chapter and provides answers to the research questions presented therein. First, the question of whether one of the six candidate noise reduction techniques out-performs all other techniques in minimizing noise in NDVI time series while preserving the integrity of the signal of interest, is addressed in section 5.1. This is followed in section 5.2 by a discussion of those factors that demonstrate an effect on the performance of the candidate techniques, such as land cover or noise level. And finally, section 5.3 discusses the benefit of applying noise reduction to NDVI time series and the circumstances in which it could be considered advantageous or disadvantageous to apply such techniques.

5.1 Overall Performance of the Candidate Noise Reduction Techniques

5.1.1 Function-Fitting Techniques

Both the Double Logistic and Asymmetric Gaussian function-fitting techniques demonstrated overall superiority in the minimization of noise in NDVI time series and the maintenance of signal integrity. While the Double Logistic function displayed slightly better overall performance than the Asymmetric Gaussian function according to standardized summary performance scores, the difference between them was negligible. Indeed, such strong similarity between the two methods was expected given the strong resemblance in their respective implementations (Beck *et al*. 2006). It can thus be stated that no one technique out-performed all other selected techniques, but rather, that in general the two function-fitting techniques were superior to the remaining four filtering techniques (i.e. the MVI, ARMD3-ARMA5, Savitzky-Golay and 4253H-Twice filters) across the variety of conditions tested in this analysis.

The reasons for the superior performance of both the Asymmetric Gaussian and Double Logistic function-fitting approaches likely include i) their strict preservation of the upper envelope of NDVI values in a time series, and ii) their reliable approximation of a relatively constant winter NDVI that reflects that season's dormant vegetative conditions. That is, by maintaining the upper envelope of NDVI values these function-

fitting techniques were better able to deal with noise that is generally negatively biased than those techniques which do not incorporate such a bias, and by maintaining constant winter NDVI in many cases, these two techniques often resulted in more accurate metric extraction. Examples of this are found in Figure 5-1, which shows the application of the six candidate techniques to the Alpine time series with 10% introduced noise, and the subsequent extraction of the start of season (SOS) and end of season (EOS) metrics. Both the Asymmetric Gaussian and Double Logistic function-fitting techniques show very effective maintenance of upper NDVI values during the growing season, while the 4253H-Twice filter underestimates these values slightly, the MVI and ARMD3-ARMA5 filters underestimate these values noticeably and the Savitzky-Golay filter tends to overestimate them, particularly in over the 2003 and 2004 growing seasons (Figure 5-1). In addition, both function-fitting methods show the effective approximation of winter NDVI values, and are not affected by the sudden dips in second, 43rd and 65th composite periods as are the Savitzky-Golay filter and to some degree, the 4253H-Twice filter. While both the MVI and ARMD3-ARMA5 filters show similar winter NDVI trajectories, they are not able to approximate the suddenness of the beginning and end of the three growing seasons as are the Asymmetric Gaussian and Double Logistic methods (Figure 5-1). Not only are the extraction of maximum NDVI, NDVI amplitude and other related metrics likely to be generally more accurate with the application of the two function-fitting techniques, but so are the subsequent extraction of SOS, EOS, LGS, and others.

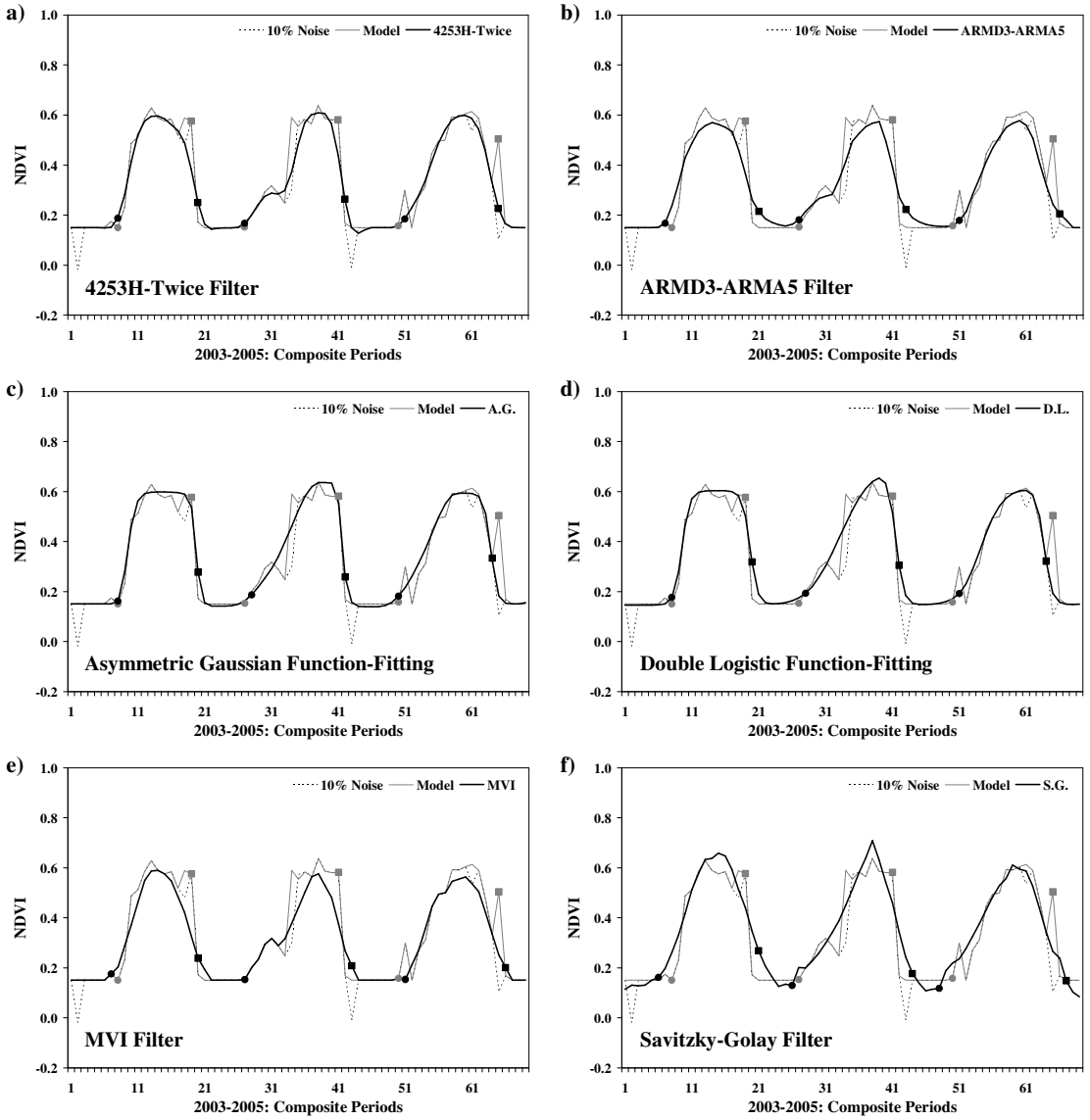

Figure 5-1: Alpine model and (70%) noisy NDVI time series, showing application of the a) 4253H-Twice filter, b) ARMD3-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter. SOS (circles) and EOS (squares) metrics derived from both the noise-reduced time series (black) and model time series (grey) are depicted.

These results reflect what is found in the current literature. The successful performance of the Asymmetric Gaussian function-fitting is supported by Jönsson and Eklundh (2002, 2004), who compared this technique with the Best Slope Index Extraction (BISE) and a Fourier-based technique, and with the Savitzky-Golay filter and a harmonic-based method, respectively. In the former, the Asymmetric Gaussian technique was first introduced and compared favourably to these other, more common

methods of noise reduction (Jönsson and Eklundh 2002). In the latter, the technique was shown to provide more flexibility than a harmonic-based technique but was also less affected by high levels of noise than was the Savitzky-Golay filtering technique, leading to conclusions of its superiority for noise minimization and the subsequent derivation of phenological metrics (Jönsson and Eklundh 2004). A comparison of the Asymmetric Gaussian and Double Logistic functions by Beck *et al*. (2006) led to similar conclusions about their relative performances as are drawn here. While both were found to perform better than a Fourier-based technique and similar to one another, the Double Logistic technique demonstrated slightly better performance than the Asymmetric Gaussian technique. Beck *et al*. (2007) further demonstrated the successful application of the Double Logistic function-fitting technique in their creation of a new data set for studying vegetation phenology and climatic factors over northern Europe.

5.1.2 Filtering Techniques

While the Double Logistic and Asymmetric Gaussian function-fitting techniques showed the overall best performance, the MVI and ARMD3-ARMA5 filters generally showed the overall poorest across the standardized summary, RMSE and metric performance scores (see Figures 4-1 and 4-2). This poor overall performance by both filters is likely the result of i) their lack of a bias toward higher NDVI values, unlike the Savitzky-Golay filter and Double Logistic and Asymmetric Gaussian function-fitting techniques, and ii) their considerable reliance on averaging techniques. By not incorporating the preferential preservation of higher NDVI values, both the MVI and ARMD3-ARMA5 filters are likely to be more greatly affected by the negatively-biased noise found in NDVI time series, and would not therefore provide as accurate approximations of original NDVI time series, leading to erroneous derivations of maximum NDVI, NDVI amplitude, and related metrics. In addition, as van Dijk *et al*. (1987) and Jönsson and Eklundh (2006) observed, running means tend to alter the shape and amplitude of the original data set, even when they incorporate weighting factors in their implementation. In other words, because such techniques simply average all noise in the time series, the calculations can be dominated by outliers, which will extend their undesirable influence to all neighbouring data points within the averaging window

(Velleman 1980). Thus, by involving the extensive use of running means the MVI and ARMD3-ARMA5 filters are more likely to distort sudden but relevant rises or falls in the NDVI time series that are phenologically significant, such as a rapid spring green-up or sharp autumnal senescence. This is illustrated in Figure 5-1. Not only were both filters unable to capture the full amplitude of the three Alpine growing seasons in this case, but neither was able to reconstruct the sharp drop and sudden end of the growing season, particularly for the 2003 and 2004 seasons (Figures 5-1b and 5-1e).

The generally poor performance of the MVI and ARMD3-ARMA5 filters found in the current research contradicts what is found in the literature. Both of these filters demonstrated the overall least ability to minimize noise and preserve data integrity in the current evaluation. However, Ma and Veroustraete (2006) concluded from their comparison of the MVI filter, the BISE technique and a Fourier transform-based technique that the former provides an effective method for minimizing noise in NDVI time series. Not only did the MVI filter demonstrate minimal distortion from the original data points in contrast to the Fourier-based method, but was better able to identify and remove spurious highs than the BISE algorithm. Filipova-Racheva and Hall-Beyer (unpublished) likewise maintained the superiority of the ARMD3-ARMA5 filter over a set of alternative running mean, running median and running maximum filters. They argued for the ability of the ARMD3-ARMA5 filter to remove spurious dips and spikes from a noisy NDVI time series while maintaining the basic shape and amplitude of the original data. It must be recognized, however that the evaluations undertaken by Ma and Veroustraete (2006) and Filipova-Racheva and Hall-Beyer (unpublished) were primarily visual and qualitative in nature. The more quantitative nature of the present evaluations and their incorporation of the additional techniques not evaluated by the above authors, may easily account for the disparity between the present and past findings.

Unlike the Double Logistic and Asymmetric Gaussian function-fitting techniques and the ARMD3-ARMA5 and MVI filters, neither the 4253H-Twice filter nor the Savitky-Golay filter demonstrated consistency in their performance across the standardized summary, RMSE and metric performance scores (see Figures 4-1 and 4-2). The 4253H-Twice filter showed better performance in terms of metric results than RMSE

results while the Savitzky-Golay filter, showed better performance than any techniques according to RMSE results, but poorer performance according to metric results. This disparity is no doubt connected not only to differences between the two filters, but to differences between the RMSE and metric-based methods of evaluation. RMSE is a measure of the average difference between two sets of values, calculated per observation or data value (Equation 3-9), which in the current case correspond to original model and noise-reduced NDVI time series values. RMSE therefore represents a general assessment of how well the application of a noise reduction technique returns a noisy time series to its original model state. The metric portion of the present evaluation, however, reflects the ability of a noise reduction technique to maintain the integrity of the original model times series when applied to one with introduced noise; it is an indirect evaluation of how well the amplitude and shape of the original time series is preserved because the derivation of time series metrics will change when the characteristics of the time series are modified.

The results indicate, then, that while the Savitzky-Golay filter generally produces noise-reduced time series that show the least amount of mean difference with the original and can therefore be said to provide the best noise minimization as indicated by RMSE results, it cannot preserve the shape and integrity of the original data as well as most other candidate noise reduction techniques, according to metric results. The 4253H-Twice filter, on the other hand, preserves the integrity of the original time series as well as both function-fitting approaches, but results in an overall greater deviation between the noise-reduced and original time series. The reasons for this likely involve i) the reliance of the 4253H-Twice filter on running median filters, and ii) the tendency of the Savitzky-Golay to preserve the upper envelope of NDVI values in a time series, a bias that is not inherent to the 4253H-Twice filtering process. That is, the incorporation by 4253H-Twice filter of a number of running medians would lead to its greater ability for maintaining the shape of an NDVI time series because, as explained by Rabiner *et al*. (1975), running medians are able to remove sharp discontinuities in a data set and follow polynomial signals, and therefore, are able to preserve the quadratic shape of an NDVI profile. Indeed, Velleman (1980) and van Dijk *et al*. (1987) all praised the superior

performance of the 4253H-Twice filter regarding its performance in comparisons with alternative nonlinear filters and Fourier-based, polynomial and running mean techniques, respectively. Its ability to handle complex time series while eliminating spurious drops or spikes was lauded in both sets of studies. However, the current evaluation, though showing the 4253H-Twice filter to perform fairly well overall, also demonstrated a greater ability by the Double Logistic, Asymmetric Gaussian and Savitzky-Golay techniques to minimize overall noise. As suggested above, this is likely because the 4253H-Twice filter does not incorporate a bias toward higher NDVI values as do these other techniques. An example of this is illustrated in Figure 5-2, which shows the application of the Savitzky-Golay and 4253H-Twice filters to the Boreal time series with 70% introduced noise. The advantage of preserving the upper envelope of NDVI values is evident, and here leads to a lower (i.e. better) RMSE for the Savitzky-Golay-filtered time series (Table 5-1). However, the timing of both SOS and EOS for the 2004 and 2005 growing seasons are more accurately derived after application of the 4253H-Twice filter; the Savitzky-Golay filter does not capture the suddenness of the beginnings and ends of the growing season as well as the 4253H-Twice filter (Figure 5-2).

The tendency of the Savitzky-Golay filter to preserve the upper envelope of NDVI values is a likely reason for its superior ability to minimize noise in general, as demonstrated by its low overall RMSE performance score. However, its inability to preserve overall shape in a time series as suggested by the metric results is perhaps, at least in part, a function of user error in selecting the window widths for the three fitting steps implemented in TIMESAT (see Chapter Three). For instance, if too wide a window is chosen the filter is not able to follow rapid changes in the time series, while too small a window would over-fit the noise-reduced time series and retain too much of the noise (Chen *et al*. 2004, Jönsson and Eklundh 2006). The Savitzky-Golay-filtered time series presented in Figures 5-1 and 5-2 suggest that the selected window widths were too wide to enable the filter to capture the sudden spring green-up and sharp end-of-season senescence. In addition, the winter NDVI profiles in the Savitzky-Golay time series display a noticeable valley-like appearance, with NDVI decreasing toward a mid-winter point and then beginning immediately to increase toward the start of the next growing

season. This can also be observed in Figures 5-1 and 5-2, and leads to a tendency to underestimate SOS and overestimate EOS, which then results in inaccurate derivation of LGS and other time series metrics that rely on these metrics for their own calculation.

Table 5-1: RMSE results calculated from the candidate noise reduction techniques.

	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	None
Grassland							
10	0.046	0.060	0.066	0.062	0.052	0.066	0.092
40	0.128	0.143	0.085	0.100	0.153	0.084	0.185
70	0.240	0.266	0.170	0.168	0.269	0.173	0.265
Parkland							
10	0.035	0.050	0.050	0.052	0.049	0.057	0.027
40	0.048	0.062	0.067	0.069	0.071	0.067	0.083
70	0.062	0.074	0.076	0.081	0.082	0.071	0.079
Boreal							
10	0.046	0.066	0.053	0.056	0.057	0.068	0.094
40	0.088	0.099	0.095	0.096	0.089	0.093	0.159
70	0.177	0.176	0.131	0.137	0.189	0.126	0.202
Lower Foothills							
10	0.097	0.093	0.086	0.082	0.080	0.084	0.149
40	0.192	0.175	0.136	0.118	0.240	0.106	0.265
70	0.272	0.270	0.182	0.179	0.244	0.140	0.320
Montane							
10	0.056	0.069	0.085	0.088	0.073	0.089	0.059
40	0.064	0.073	0.131	0.130	0.074	0.140	0.123
70	0.095	0.093	0.147	0.147	0.088	0.152	0.170
Alpine							
10	0.061	0.069	0.057	0.060	0.074	0.072	0.068
40	0.137	0.130	0.111	0.102	0.120	0.111	0.158
70	0.166	0.150	0.116	0.118	0.161	0.112	0.205

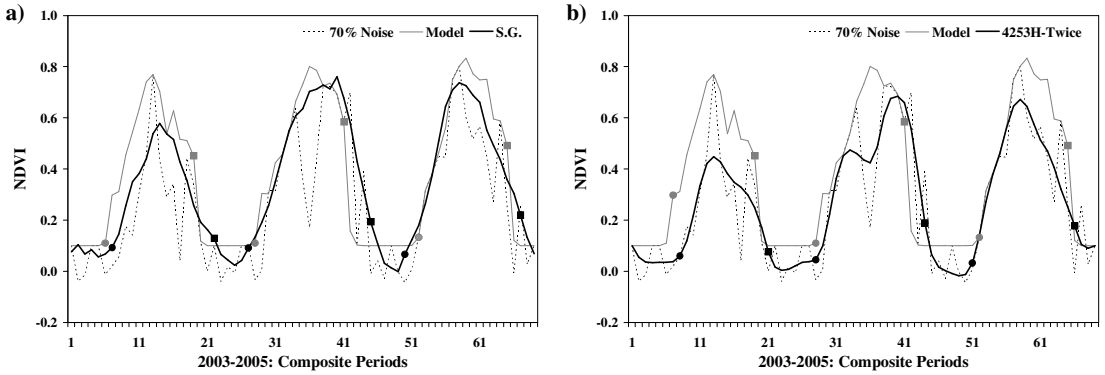

Figure 5-2: Boreal model and (70%) noisy NDVI time series, showing application of the a) Savitzky-Golay and b) 4253H-Twice filters. SOS (circles) and EOS (squares) metrics derived from both the noise-reduced time series (black) and model time series (grey) are depicted.

Chen *et al*. (2004) demonstrated the superiority of the Savitzky-Golay filter over Fourier-based techniques, and a comparable performance to the BISE method, but Jönsson and Eklundh (2004), however, express difficulties in deriving seasonal parameters or time series metrics from time series on which the Savitzky-Golay filter was applied, remarking that the filter was more affected by greater amounts of noise than was the Asymmetric Gaussian function-fitting technique. Indeed, one can observe how rough the Savitzky-Golay time series appears when compared to the smooth appearance of both the Double Logistic and Asymmetric Gaussian functions (Figure 5-1). It is quite likely that this high-frequency variability could also contribute to less accurate metric derivations.

According to standardized summary, RMSE and metric performance scores the very worst overall performance was produced by not applying noise reduction. This was anticipated, given that not only would one expect noisy time series to generally show the greatest mean deviation from the original model time series, but one would expect the least accurate metric calculations to be made from the noisy time series on which no attempt at rectification had been made. One would conclude from this that, overall, any noise reduction is generally more favourable than nothing at all. However, a closer examination of these results reveals individual scenarios where not applying noise reduction produced the same or better results than its application, suggesting that general

observations may not apply to all cases. These contradictory results are discussed further in section 5.3 below.

5.2 Factors Influencing Performance

5.2.1 Land Cover

Some of the principal observations made above regarding the overall performance of the candidate noise reduction strategies, were also found to exist across the six land cover types represented in the present research. In general the two function-fitting procedures (Double Logistic and Asymmetric Gaussian) consistently out-performed both the MVI and ARMA3-ARMD5 filters, and across the majority of land covers the unfiltered noisy data performed the worst. Some consistency in particular trends thus exists across land covers, suggesting some robustness in these results. Despite these general trends, however, considerable variation in the performance of the candidate noise reduction techniques was also evident with the stratification of the results by land cover.

One would expect that the observed influence of land cover on the performance of the six candidate noise reduction techniques would reflect the effects of varying phenological patterns and the differing abilities of the candidate techniques to approximate particular types of time series profiles. This was observed to some degree with the Alpine land cover, which showed a plateau-like growing season curve that was generally better approximated by the Asymmetric Gaussian and Double Logistic fitting function techniques than alternative techniques (Figure 5-1). However, the remaining land cover model time series showed very similar profiles to one another with only minor differences between them, as can be seen in Figure 3-4 (Chapter 3). All but the Alpine land cover showed a sudden onset of spring at the start of the growing season, a rapid increase toward a narrow peak in growing season greenness and an immediate commencement of rapid senescence and an abrupt end to the growing season. This similarity between all but the Alpine model time series suggests that much of the observed variation in noise reduction technique performance with land cover is not likely to be the result of differing phenological patterns across land cover types. Rather, other factors are likely the cause of this variation.

Closer examination reveals that the strength and type of noise introduced from the noisy pixel drawn from each of the land cover types was the most probable cause for the varying performances of the noise reduction techniques over the six land cover types. That is, different land cover time series with the same percentage of introduced noise can show very different patterns in terms of the strength and direction of noise that is present, and this produces varying performance over the land cover types by each of the noise reduction techniques. For instance, Figure 5-3 shows the Grassland, Parkland and Montane time series with 70% noise, with the 4253H-Twice filter and Double Logistic function-fitting technique applied to each. Although 70% of the dates in all six time series have been replaced with dates from a noisy pixel, this noise varies in its strength and character between the three land cover types.

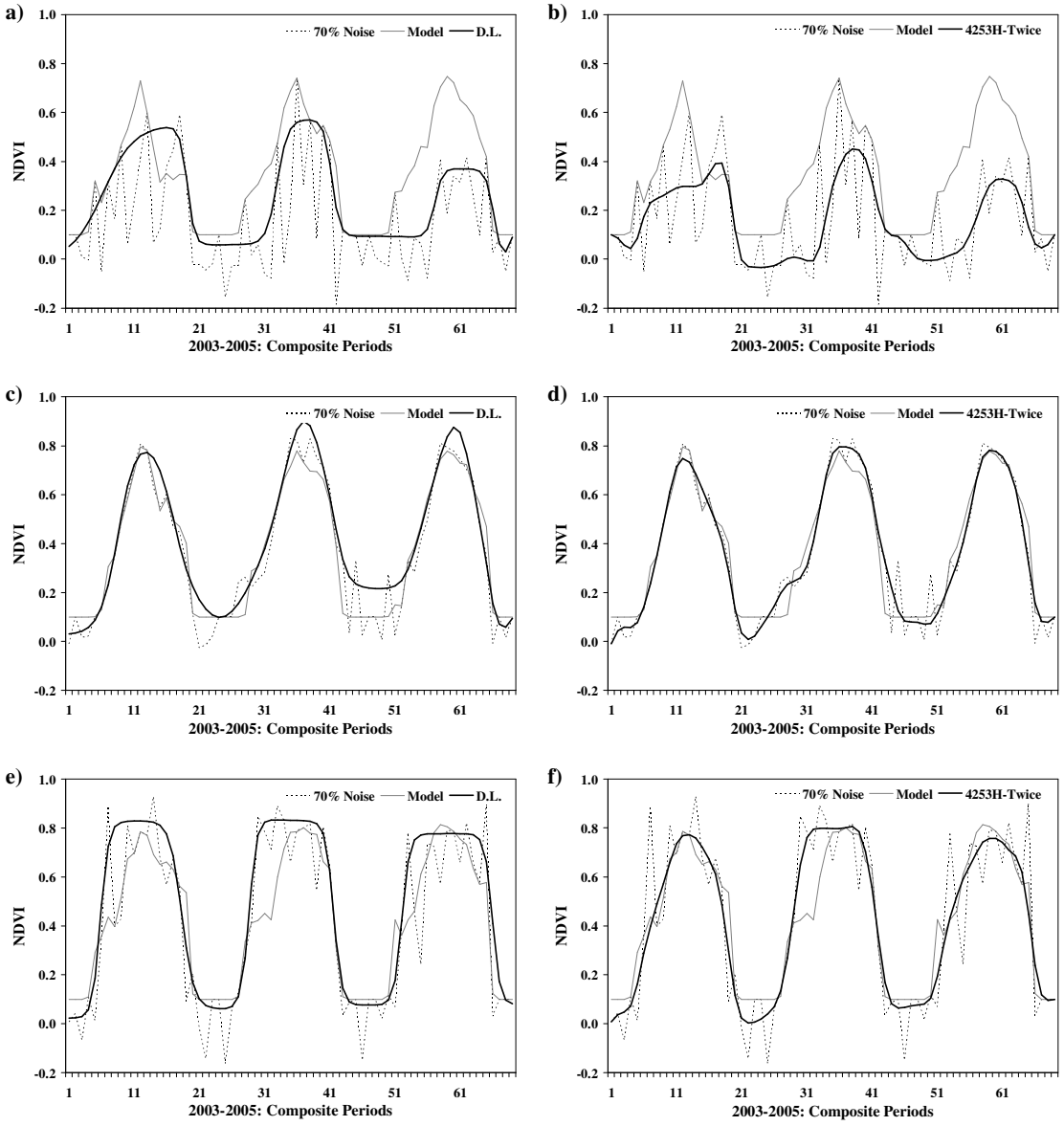

Figure 5-3: Grassland (a, b), Parkland (c, d) and Montane (e, f) model and (70%) noisy NDVI time series, showing application of the Double Logistic function-fitting technique (left) and the 4253H-Twice filter (right).

Both the Grassland and Montane time series contained considerably stronger noise than the Parkland time series, which showed remarkably little deviation from the original, given the 70% noise level (Figure 5-3). In addition, while the noise in the Grassland time series is very negatively biased, the noise introduced to the Montane time series shows a notably greater amount of spurious highs, particularly during the growing season. The Parkland time series also displays both false highs and lows in the

introduced noise, though its overall noise is considerably weaker (Figure 5-3). This variation in noise between land cover types affects the performance of noise reduction strategies, as the comparison between the 4253H-Twice filter and Double Logistic technique in Figure 5-3 demonstrates. Here, the positive noise in the Montane land cover led to the overestimation of growing season NDVI by the Double Logistic function-fitting, where the 4253H-Twice filter for the most part was much better able to cope with the spurious highs in the time series. A similar result is seen for the Parkland time series, where positive errors in NDVI led to differing performances by the two techniques (Figure 5-3). However, the preservation of the upper envelope of NDVI values by the Double Logistic method led to better approximations of growing season amplitude and winter NDVI for the Grassland land cover type, while the 4253H-Twice filter was more greatly affected by the negative noise in this noisy land cover time series (Figure 5-3). Such variations in the performances of the Double Logistic function-fitting technique and the 4253H-Twice filter were not the result of variations in phenology between the three land cover types, but rather the strength and nature of the noise introduced to the time series. One can thus conclude that in the current study, land cover phenology shows little influence on the ability of the six candidate noise reduction techniques to minimize noise and the integrity of the NDVI signal. Rather, the strength and character of the noise present in these time series is the cause of variation in performance between land cover types. On the basis of these findings, it is advisable to consider the strength and potential biases that may be present in a multi-temporal NDVI data set of interest while selecting an approach to noise reduction.

Few discussions concerning the effect of individual factors on the application of the candidate noise reduction techniques are found in the literature, though some comparison of a few noise reduction techniques applied across varying vegetation types and land cover was undertaken by both Chen *et al*. (2004) and Ma and Veroustraete (2006). In the former, the Savitzky-Golay filter was compared with a Fourier-based technique and the BISE algorithm over a variety of test pixels covering various land covers across continental China. Particular pixels were used to illustrate the poor performance of the Fourier-based methods over particularly asymmetric phenological

profiles (i.e. NDVI time series), including evergreen forest and double-cropping rice fields, but further discussion regarding the effects of land cover on the performance of the three techniques was not presented (Chen *et al*. 2004). Ma and Veroustraete (2006) also applied three noise reduction techniques to a number of test pixels covering various land cover types, but again did not discuss the effects of these differing vegetative communities on the performances of the MVI, BISE and Fourier-based techniques. The current analysis appears unique in its investigation of the potential effects that land cover or other factors may have on the reduction of noise in NDVI time series.

5.2.2 Noise Level

As discussed above, the strength and type of noise introduced to each of the model NDVI time series caused considerable variation in the performance of the candidate noise reduction techniques across the six land cover types. The stratification of standardized performance scores by noise level (i.e. the percentage of introduced noise), further demonstrated the influential nature of noise; while some techniques performed better with less noise, others improved with increasing levels of noise, and no one technique performed best across all three noise levels. One interesting, yet not unexpected trend that deserves discussion is the relationship between performances of the six noise reduction strategies and the application of no noise reduction over the three levels of introduced noise. At both the 40% and 70% noise levels, not applying any noise reduction technique produced the worst overall standardized summary performance score. However, at the 10% noise level all candidate techniques, with the exception of the 4253H-Twice filter, performed similar to or worse than no noise reduction (Figure 4-6). The Savitzky-Golay filter performed particularly poorly at this lowest noise level. This general trend is not surprising but is important in its suggestion that the application of a noise reduction technique to a clean time series will tend to degrade it. That is, with only 10% or seven of the 69 data values in the three-year time series being replaced with noisy NDVI values, and with the candidate noise reduction techniques operating on the entire time series regardless of the number of noisy dates, the ‘noisy’ time series will be closer to the original model than one on which noise reduction has been applied. When greater levels of noise are encountered, such as 40% and 70% time series, any noise

reduction technique will in general show a better standardized performance score than if no noise reduction were applied. This is clearly demonstrated in Figures 5-1, 5-2 and 5-3. For example, the former shows that at the 10% noise level the peculiar shape and irregularity of the original Alpine model time series is most closely followed by the unfiltered, noisy time series; none of the candidate noise reduction strategies are able to replicate the original time series without causing more alteration than was caused by the introduction of a few noisy dates (Figure 5-1). Both Figures 5-2 and 5-3, however, which show the Boreal, Grassland, Parkland and Montane time series with 70% noise, illustrate how the application of any of the noise reduction techniques presented here will tend to reduce the high level of noise and return these noisy time series to a state closer to the original model time series than is otherwise found in the noisy time series (Figures 5-2 and 5-3). This is particularly evident in the case of the Grassland time series.

A second interesting but rather unexpected trend observed in the variation of performance with noise was that not applying noise reduction produced a notably higher (i.e. worse) standardized summary performance score at the 40% noise level than at the 70% noise level (Figure 4-6). This would seem to indicate that overall, noisy time series with 40% noise are worse than those with 70% noise. However, closer examination of the data reveals this to be an artefact of the standardizing procedure applied to the performance scores. Inspecting the raw, *unstandardized* performance scores for each candidate technique presented in Tables 5-2 and 5-3, reveals that not applying noise reduction actually does produce a higher (i.e. worse) total performance score at the 70% noise level than the 40% noise level. The standardization procedure involved dividing the total performance scores for each noise reduction technique by the mean total performance score of all techniques for that particular noise level. This resulted in standardized RMSE and metric performance scores that were then summed to produce standardized summary scores (see Chapter Three for full description). The higher (i.e. worse) standardized performance scores produced at the 40% noise level for time series on which no noise reduction was applied indicates that the total performance scores for this level of noise were much higher in comparison to the mean performance, than what is observed at the 70% noise level. In other words, noisy time series perform worse in

comparison to a mean performance at the moderate noise level, than what is seen at a higher noise level. This would suggest that the application of noise reduction generally produces better noise minimization and preservation of data integrity when compared to no application of noise reduction at moderate noise levels than at high noise levels. This further emphasizes the necessity of assessing the strength, level and nature of noise present in a data set of interest when addressing the issue of noise reduction and NDVI time series; users are ill-advised to neglect a consideration of these factors when employing these data sets in the remote sensing of vegetation phenology.

Table 5-2: Unstandardized RMSE performance scores, stratified by noise level.

Noise Level	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	None
10%	35	41	40	40	38	51	42
40%	66	68	64	62	74	60	96
70%	102	103	83	84	103	77	123

Table 5-3: Unstandardized metric performance scores, stratified by noise level.

Noise Level	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	None
10%	286	357	335	316	360	448	318
40%	352	418	364	350	423	432	682
70%	455	513	389	407	535	466	666

5.2.3 Time Series Metric

In addition to demonstrating a variation in performance with land cover, noise level, and type and strength of noise, the six candidate noise reduction techniques also showed considerable variation in performance when evaluated by metric. Indeed, little consistency was observed in the relative performances of the candidate techniques across the ten metrics (Table 4-8). One of the few patterns that can be discerned is that of four of the five metrics directly related to the timing of phenological processes – SOS, EOS, LGS and timing of maximum green-up – all demonstrated the worst standardized summary performance scores from time series on which no noise reduction has been

applied (Table 4-8). In other words, it appears that the application of any noise reduction technique will generally produce a better performance score for these four metrics than when noise reduction has not been applied. This is likely a product of the high-frequency perturbations in the noisy time series. While no strategies result in a complete removal of noise, nor a recovery of the original model time series, it is important to recognize that they all serve to reduce high-frequency noise and, at the very least, reduce the probability that a sudden spike or drop could be mistaken for the timing of an event such as SOS or EOS. Examples of this are found in Figure 5-4, where the derivation of SOS and EOS from noise-reduced, original model and noisy time series for the Boreal and Lower Foothills land covers at 40% and 70% noise, with application of the 4253H-Twice filter, are presented.

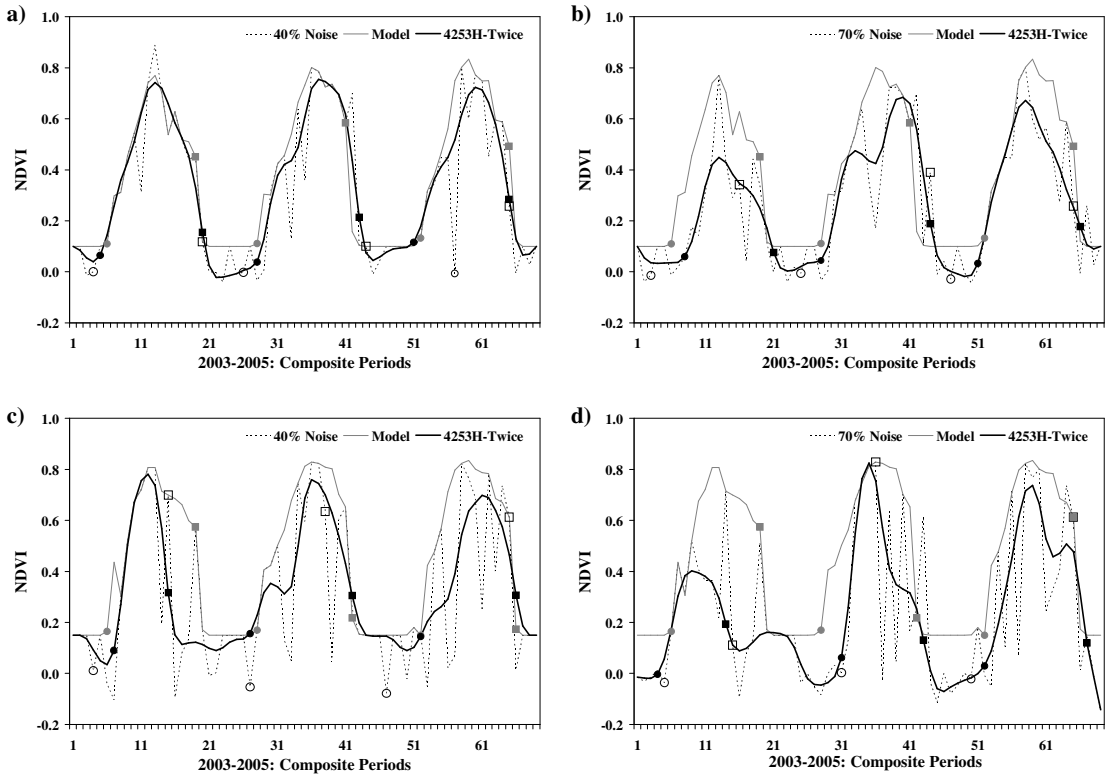

Figure 5-4: Boreal (a, b) and Lower Foothills (c, d) model and (70%, 40% respectively) noisy NDVI time series, showing application of the 4253H-Twice filter. SOS (circles) and EOS (squares) metrics derived from the noise-reduced time series (black), model time series (grey) and noisy time series (white) are depicted.

In Figure 5-4, the influence of severe irregularity on the derivation of time-based metrics from noisy time series is evident, particularly at the beginning of the 2005 and

2004 seasons. In these cases, mid-winter drops that are followed by a spike are mistaken for early SOS dates (Figure 5-4). In addition, a number of the EOS dates, particularly 2004 EOS for the Lower Foothills time series, show early estimations as derived from the noisy time series because of late season drops that occur prior to the true end of the growing season (Figure 5-4). It is also clear from this figure how easily the timing of maximum NDVI and the timing of maximum green-up could be inaccurately derived. Oftentimes, the dates at which maximum NDVI occurs in the original model time series have been replaced by noisy values during the noise introduction process, and thus no longer represent maximum NDVI. In addition, the sharp slopes created by the juxtaposition of drops and spikes in the noisy time series could lead to inaccurate identifications of the timing of maximum rates of green up (Figure 5-4). It is thus not very surprising that the application of noise reduction would have a positive effect on subsequent derivations of these time-based metrics from NDVI time series. One may therefore argue that if the derivation of metric related to the timing of the growing season and vegetative development is of particular interest, the application of noise reduction to NDVI time series should be considered. However, if alternative metrics unrelated to timing, such as the maximum level of photosynthesis reached during the growing season (i.e. maximum NDVI), the application of noise reduction would not be recommended on the basis of these findings.

5.2.4 Year

In addition to stratifying the candidate noise reduction techniques on the basis of land cover, noise and metric, a comparison of these techniques and their performance between the three years comprising the data set also revealed some variation in the results of the evaluation. The results of this analysis, however, showed few annual differences, which suggest an element of robustness in the relationship between the varying performances of the six techniques with time. This is supported by the observation that less variation in standardized performance scores exists between the three years for each noise reduction technique than between techniques for each year. In addition, reminiscent of overall general trends discussed above, the Double Logistic, Asymmetric

Gaussian and 4253H-Twice noise reduction techniques generally outperformed the MVI, ARMD3-ARMA5 and Savitzky-Golay filters across the three years.

What variations did exist in technique performance across the three years were likely the result of differences in noise over the three years, in addition to slight inter-annual variations in the phenological curves. This variation in annual noise patterns is particularly evident in the Grassland time series depicted in Figure 5-3, where noise shows an especially strong negative bias in 2005 when compared to 2003 and 2004, and limits the ability of both the 4253H-Twice filter and Double Logistic function-fitting technique to return the 2005 portion of the time series to its original model state. Similar examples are found in the second and third graphs of Figure 5-4, where the Boreal and Lower Foothills time series with 70% noise both show a greater amount of negatively biased noise in 2003 than in 2004 or 2005, which again limits the performance of the 4253H-Twice filter for this particular year.

5.3 The Case Against Noise Reduction

Several of the results of the scenario analysis presented in section 4.2 appear contradictory to the standardized performance score findings. In particular, the scenario analysis revealed that overall, in more than half of cases, the unfiltered noisy NDVI time series performed the same or better than a time series to which noise reduction had been applied. This contradicts the observation according to overall standardized performance score results that the raw unfiltered data produced the worst overall standardized results. While the latter analysis supports the use of noise reduction, the former suggests it is not generally beneficial. This disparity likely reflects differences in the evaluation techniques between these two analyses rather than conflicting results. Standardized performance scores indicate a general ability to minimize noise and preserve data integrity regardless of the ability of other techniques in any one particular case. They are standardized summations of raw performance scores from a multitude of scenarios and therefore represent something of a total amount of deviation in RMSE and metric calculations. The scenario analysis, however, denoted the number of times noise reduction performed better than the raw, noisy data. On the basis of these two sets of results it is likely that, while comparison of each noise-reduced time series to the

equivalent noisy time series typically demonstrated no benefit by the former, when the performance of the raw, unfiltered data was worse, it tended to be considerably worse, contributing to higher (i.e. worse) overall standardized performance scores. That is, in many situations noise reduction did not prove advantageous, but where it did, the benefit was substantial when compared to the alternative.

The assertion that noise reduction provides no benefit in the majority of cases does not hold true of all scenario analysis results, however. When the overall total was divided into its RMSE and metric scenario components, the former demonstrated improved performance by noise-reduced time series over noisy time series in more than half of RMSE scenarios while the latter showed the same or better performance scores by noisy time series in the majority of metric scenarios. The total RMSE scenario tallies generated results opposite to the overall results but the total metric scenario tallies produced results similar to the overall tallies. These seemingly contradictory outcomes actually reflect the different aspects of noise reduction that these two sets of scores evaluate. RMSE measures mean deviation between both the noisy and noise-reduced time series and the original model time series, while metric accuracy is a measure of how well the shape and amplitude of the original signal are preserved. When examined more closely it is, in fact, not surprising that such a contradiction between these two sets of evaluations can arise. For example, Figure 5-5 presents the application of the 4253H-Twice filter to the Montane time series with 10% noise, and the Grassland time series with 70% noise, along with the derived timings of SOS and EOS from the model, noisy and noise-reduced time series. In both cases the 4253H-Twice filter produced a lower (i.e. better) RMSE than the noisy time series (Table 5-1). However, one can also observe that in both cases there are several instances where the noisy time series produced more accurate SOS and EOS calculations. In the Montane time series the timing of SOS in all three years was either closer to the model SOS when derived from the noisy time series than when derived from the filtered time series or exactly equivalent to the model SOS, where SOS from the filtered time series was not (Figure 5-5). It is evident that the alteration in the shape and sharpness of late winter and early spring NDVI values was the cause of these inaccuracies in the time series on which the 4253H-Twice filter was

applied. The same is also true of the timing of EOS in 2004 and 2005 (Figure 5-5), particularly in the second graph. Here, the timing of SOS was much more accurately derived from the noisy unfiltered time series for both the 2004 and 2005 growing seasons, and for the 2003 and 2005 growing seasons the EOS metric derived from the noisy time series matched that derived from the model time series, while EOS derived from the filtered time series did not. Thus, while applying the 4253H-Twice filter to the Montane and Grassland time series presented in Figure 5-5 generated good RMSE results (Table 5-1), metric scores is in several instances were worse than the unfiltered time series. This means that while applying noise reduction did minimize noise and return the noisy time series to its original condition to some degree, the integrity of the time series could be so much altered by this that derived metrics were often just as accurate or more accurate when noise reduction was not applied. This dichotomy between RMSE and metric scenario results generally persists even with stratification by land cover, noise level, year and metric, indicating that the above assertion is a valuable observation which remains consistent across a variety of conditions.

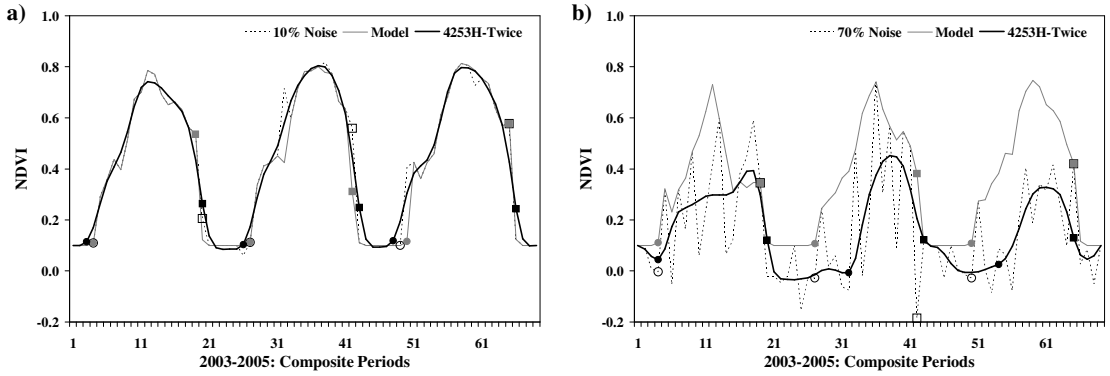

Figure 5-5: Montane (a) and Grassland (b) model and (10%, 70% respectively) noisy NDVI time series, showing application of the 4253H-Twice filter. SOS (circles) and EOS (squares) metrics derived from the noise-reduced time series (black), model time series (grey) and noisy time series (white) are depicted.

5.3.1 Influential Factors

When stratified on the basis of noise reduction technique, both RMSE and metric scenarios showed relatively little variation (Figures 4-14 and 4-17). All techniques showed that 67% to 82% of RMSE scenarios favoured the use of noise reduction, and that 56% to 68% of metric scenarios show no improvement in performance scores with

noise reduction. This consistency in numbers between the candidate techniques suggests that regardless of one's choice of noise reduction method, the application of noise reduction, while very likely to provide an improvement in terms of the overall accuracy of NDVI values, was also likely to degrade the quality of most NDVI time series metric calculations. This means that factors other than the choice of noise reduction strategy were more influential in the benefit of applying noise reduction.

Year also did not appear to strongly influence the RMSE and metric scenario results. A lack of variation with year suggests a consistency in the potential disadvantage of noise reduction with time, perhaps also indicating a similar consistency with time in the influence of additional factors. With regard to noise level, both RMSE and metric scenario tallies demonstrated a greater likelihood for the degradation of NDVI time series with the application of noise reduction at minimum levels of noise than at moderate to high levels of noise, but little variation existed across the 40% and 70% noise levels. The reason for this degradation at minimal noise levels is no doubt the result of processes similar to those discussed with regard to the effect of noise level on standardized performance score calculations – the potential for noise reduction strategies to degrade relatively clean signals. However, the lack of variability across moderate to high noise levels suggests that aside from the lowest levels of noise, which are not often encountered in NDVI time series, this factor did not strongly influence the benefit of noise reduction as it did standardized performance scores. One's choice of noise reduction technique, the year and the level of noise therefore did not affect the benefit of noise reduction to any great degree. Rather, land cover and one's choice of metric appear to have had the greatest influence on these results.

The variation in scenario tallies with land cover was most noticeable in terms of RMSE scenarios, where the percentage of RMSE scenarios favouring the application of noise reduction over no noise reduction ranged from 50% for the Parkland and Montane land cover types to 100% for the Boreal land cover type (Figure 4-7). As with the variation in technique performance over the six land cover types, this pattern is likely a result of the varying types and strengths of noise that were introduced in each. For instance, the Parkland time series tended to be degraded by the application of noise

reduction much worse than the Grassland time series at the same noise level, resulting in a greater likelihood that the application a noise reduction strategy to the Grassland series would be beneficial (Figure 5-3). In other words, because the noisy Parkland time series showed remarkable similarity to the original model time series, even when 70% of its 'clean' values were replaced, noise reduction did little to improve things. On the other hand, because the noise introduced to the Grassland time series was so degrading, the application of any noise reduction technique was an improvement. While it is likely that this is the cause of variation in the benefit of noise reduction across land cover types, it is also important to recognize that only 18 RMSE scenarios were tested, and that differences in the percentages between land cover types which seem very large, may represent a difference in only two or three scenarios in the tallied numbers (e.g. Figure 4-7). Larger sample sizes would be required to confirm these findings.

Observations regarding the second factor appearing to most influence the benefit of applying noise reduction, one's choice of metric, while not as distinct, were also less likely to be exaggerated because many more metric scenarios comprised the tallies. As Figure 4-12 demonstrates, the percentage of scenarios which favour the application of noise reduction ranged from just 15% for maximum NDVI, to 61% for the maximum rates of green-up. While it is difficult to discern strong trends across the metrics in these tally results, there is one quite remarkable observation that ought to be discussed. This is that of all the metric scenario tallies, only those tallies calculated for NDVI amplitude, maximum rate of green-up and average NDVI show results suggesting that the majority of scenarios demonstrate better performance scores when noise reduction is applied. The opposite is true for all other metric scenario tallies (see Figure 4-4, 4-6, 4-8, 4-10, 4-11 and 4-12).

These contrary numbers reflect the greater benefit of removing high-frequency variation in NDVI time series to derivations of NDVI amplitude, average NDVI and maximum green-up than for other metrics. That is, one could suggest that removing spurious spikes and drops is more likely to result in more accurate derivations of these three metrics than it is for the remaining seven metrics. For instance, because the calculation of NDVI amplitude relies on accurate derivations of both maximum and

minimum NDVI, the prominence of spurious drops and negative bias in the noisy time series is quite likely to lead to considerable inaccuracies in this metric. Although maximum NDVI was demonstrably more often accurately derived from the noisy time series, the presence of negatively-biased noise in noisy time series would often lead to gross underestimations of minimum NDVI. Thus, the application of any noise reduction technique which reduces the presence of high-frequency noise is likely to improve calculations of minimum NDVI, and consequently, calculations of NDVI amplitude. This is illustrated in Figure 5-6 and Table 5-4, which present the noisy, model and noise-reduced time series of the Montane land cover for 2004 with 40% noise. Here, application of the MVI filter removes both spurious spikes and drops in the time series, and results in a more accurate NDVI amplitude calculation than does the noisy time series.

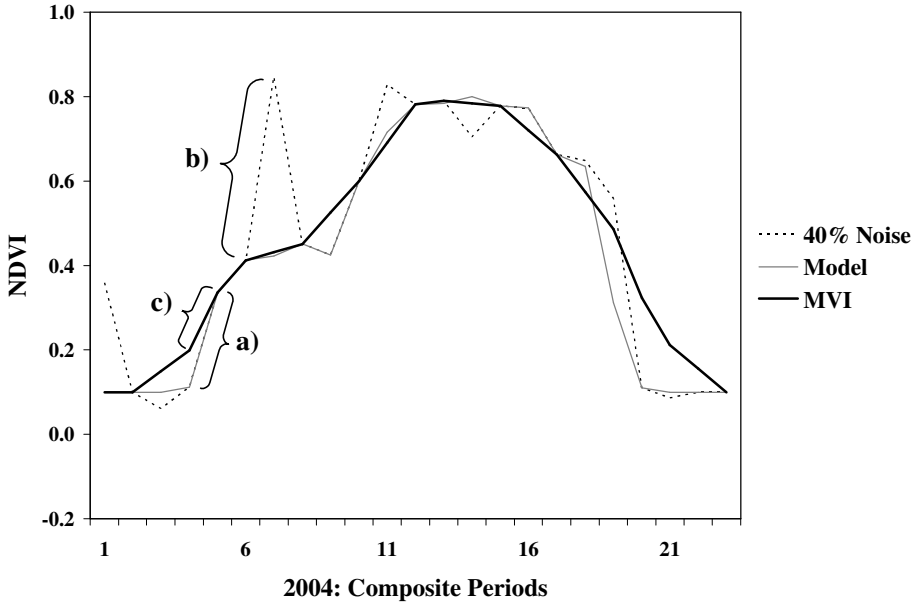

Figure 5-6: Montane model and (40%) noisy NDVI time series for 2004, showing application of the MVI filter. Location of maximum green-up is depicted for the a) original, b) noisy and c) filtered time series.

Table 5-4: Maximum green-up and NDVI amplitude calculated from the model, noisy and MVI-filtered 2004 portion of the Montane NDVI time series with 40% noise.

NDVI Time Series	Maximum Green-Up (NDVI/Composite Period)	NDVI Amplitude
Model	0.224	0.700
40% Noise	0.138	0.690
MVI-Filtered	0.435	0.774

A similar argument may be made with reference to the number of scenarios favouring the application of noise reduction for the derivation of average NDVI. The negative bias observed in the noisy time series is likely to reduce estimations of growing season NDVI, and if the application of a noise reduction technique minimizes the strength and influence of high-frequency perturbations on the NDVI times series, it follows that the subsequent calculations of average NDVI will more closely approximate the true model calculations than do those derived from the noisy time series.

The greater number of metric scenarios that produce more accurate calculations of maximum green-up with the application of noise reduction is also likely related to the removal or reduction of high-frequency noise in NDVI time series by these techniques. In the case of this metric, however, it is the presence of sudden rises and falls in NDVI that occur with the juxtaposition of extremely low and high NDVI values, which could lead to inaccurate calculations of maximum green-up. In other words, these sharp, spurious increases could easily be mistaken for high rates of green-up during spring in noisy NDVI time series. With the application of noise reduction, however, these extremes in NDVI values and high-frequency noise in the times series are generally minimized and sudden rises in NDVI are often preserved only where relevant (i.e. at the beginning of spring). An example of this can be found in Figure 5-6 and Table 5-4, where maximum green-up is more accurately derived from the MVI-filtered NDVI time series than from the noisy time series, which shows a sudden spurious spike in the early growing season that is mistakenly identified as the timing and rate of maximum green-up.

The MVI filter removes this spurious spike, and produces a smoother time series from which slightly underestimated but more accurate maximum green-up and NDVI amplitude are derived.

Despite the rationality of the above arguments, when observations regarding the metric-based stratification of standardized performance scores and the similar stratification of scenario tally results are compared, something of a contradiction arises. As discussed above, standardized performance scores indicated that for four of five time-based metrics (SOS, EOS, LGS, and time of maximum green-up) any noise reduction technique performed better than the raw, unfiltered data. However, scenario results suggest that for only three value-based metrics (NDVI amplitude, maximum green-up and average NDVI) did the application of noise reduction produce improved performance over noisy time series in most cases. An explanation such as the one given above for a similar discrepancy in overall results between performance scores and scenario tallies could likely be attributed to these seemingly contradictory outcomes. That is, the extraction of the four time-based metrics from noisy, unfiltered time series is generally more accurate than from noise-reduced time series, but where the derivation of these metrics from noisy time series is less accurate, it is considerably less so. On the other hand, applying noise reduction before derivation of the three value-based metrics is generally beneficial, but where it is not beneficial the resulting metrics are especially inaccurate in comparison to those derived from noisy, unfiltered time series. It was thus generally advantageous to reduce noise in NDVI time series before deriving these three value-based metrics but the application of such techniques often led to greater inaccuracy in the derivation all other metrics, including the four time-based measures. Nevertheless, the inaccuracy attributed to extracting SOS, EOS, LGS and time of maximum green-up from noise-reduced time series was not as great as that encountered from noisy time series in situations where noise reduction proved more beneficial. The complexity of these findings recommends caution to all users when applying noise reduction with the objective of phenological variable extraction. Such techniques very often hindered the derivation of these metrics, but neglecting to apply noise reduction also led to more serious errors in particular circumstances.

5.4 Chapter Summary

The current empirical comparison showed that over the broad range of scenarios investigated in this experiment, both of the function-fitting techniques, the Double Logistic and Asymmetric Gaussian functions, demonstrated a superior ability over the four filtering methods for minimizing noise in NDVI time series and maintaining the integrity of the original, desired signal. This superior performance is likely the result of i) a better capability for approximating constant winter NDVI, which reflects dormant vegetative conditions, and ii) a strict preservation of the upper envelope of NDVI values. Not only were they better able to deal with noise that was primarily negatively biased, but subsequent metric calculations generally tended to be more accurate because of better approximations of winter NDVI, and beginning and end of season transitions. These findings support those of Jönsson and Eklundh (2002, 2004) and Beck *et al*. (2006, 2007), who demonstrated the flexibility of these methods when confronted with irregular, asymmetrical NDVI time series, particularly in comparison with more common Fourier-based strategies. They also found these techniques to be less vulnerable to high levels of noise than other techniques, such as the Savitzky-Golay filter and BISE method.

Both the MVI and ARMD3-ARMA5 filters showed consistently poor performance in the current empirical comparison despite successful applications asserted in the literature. Both lack the ability to maintain the upper envelope of NDVI values, and rely heavily on averaging techniques. The former factor results in a greater vulnerability to negatively-biased noise; the latter, a tendency to distort the shape and amplitude of NDVI time series signals.

Both the 4253H-Twice and Savitzky-Golay filters demonstrated overall better performance than the MVI and ARMD3-ARMA5 filters, but varied in their performances according to RMSE and metric results. The Savitzky-Golay filter provided the best general noise minimization but was less capable of maintaining the integrity of original NDVI signals for subsequent metric extraction. The 4253H-Twice filter, however, was more successful at metric extraction but less so at general noise minimization. The ability of the former filter to maintain the upper envelope of NDVI values and the lack of such a bias in the latter is a likely explanation for the better general noise minimization

observed in the former. The greater capability of the 4253H-Twice filter for deriving accurate time series metrics could doubtless be explained by the reliance of the 4253H-Twice filter on running medians and the potential for user error in selecting moving window widths for the Savitzky-Golay filter.

The current research demonstrated notable variation in performance of the candidate noise reduction techniques with land cover type. This variation was interpreted not to be caused by differing phenologies across land covers as one might expect, but rather, differences in the strength and type of noise. In particular, the presence of negatively-biased noise led to underestimation of original NDVI signals by techniques which did not aim to preserve the upper envelope of NDVI values, while techniques incorporating this aim over-estimated time series for which noise was not negatively-biased. Observations of the results stratified by noise level revealed that any of the six candidate noise reduction techniques were at least somewhat successful at the higher noise levels (i.e. 40% and 70% introduced noise), while at the 10% noise level only the 4253H-Twice filter produced a better standardized performance score than no noise reduction. The application of noise reduction to time series with minimal levels of noise was found to degrade these relatively clean signals, while at moderate to high levels of noise the application of such strategies appeared advantageous.

The performance of the six candidate noise reduction strategies, though varying with land cover and strength, type and level of noise, appeared to be most affected by one's choice of metric. Indeed, the variation in noise reduction technique performance with metric was so great that few broad trends could be observed. Nonetheless, it can be argued that the application of noise reduction techniques to NDVI time series generally resulted in more accurate derivations of time-based metrics; four of the five metrics directly related to the timing of phenological processes demonstrated the worst standardized summary performance scores where no noise reduction had been applied. It was demonstrated that this likely results from an increased difficulty in deriving such time-based metrics from noisy time series containing a number of spurious spikes or drops, which could be mistakenly identified as SOS, EOS or the timing of maximum green-up. Because the application of any one of the six candidate noise reduction

techniques minimized the amount of high-frequency noise found in NDVI time series, it therefore follows that the difficulties in deriving time-based metrics from jagged, noisy time series were also minimized with the application of noise reduction.

When compared with these standardized performance score results, however, the results of the scenario analysis appeared to be contradictory. While the former demonstrated a notable support of the candidate noise reduction techniques as a superior choice to the use of uncorrected, noisy data, the latter suggested something quite different. Overall results of the scenario analysis indicated that in nearly 70% of individual RMSE and metric scenarios, the noisy, unfiltered NDVI time series performed equally to or better than a noise-reduced time series. This disparity can be attributed to differences between the two methods of analysis. While the scenario analysis represents the performance of the candidate noise reduction techniques relative to noisy data over individual scenarios, the standardized performance scores are total measures of overall inaccuracy. Thus, while noise reduction may not have produced benefit over the majority of all scenarios, for those cases where it was beneficial the alternative choice of not applying noise reduction caused considerably worse errors than applying noise reduction when it was not beneficial.

The same was true of all metric scenarios, when examined separately, but RMSE scenarios showed an opposing trend, demonstrating better performance by a noise reduction technique in more than half of RMSE scenarios. Thus, while the application of noise reduction was frequently successful in terms of the general minimization of noise, it generally led to less accurate metric derivations because the NDVI signal could be considerably altered with the application of such techniques.

Stratification revealed little variation in these results with either the selection of noise reduction technique, noise level or year; none of these factors demonstrated a strong effect on the probable benefit of applying noise reduction. Stratification by land cover generated considerable variation in scenario analysis results however, particularly in the case of RMSE scenarios where half to 100% of scenarios demonstrated improved performance with the application of noise reduction across the six land cover types. Nonetheless, the small number of scenarios on which these numbers were based

necessitates the acquisition of larger sample sizes before the accuracy of these observations could be confirmed.

The factor appearing to most influence the potential benefit of noise reduction for NDVI time series is one's choice of metric. Scenario results varied considerably with metric, and are not likely to be exaggerated because metric sample sizes are much larger than for RMSE results. One of the few patterns discernable from these stratified metric scenario results is the contradictory findings observed for NDVI amplitude, average NDVI and maximum green-up. All metric scenario results indicated an equal or improved performance by unfiltered, noisy NDVI time series than by noise-reduced time series, with the exception of these three metrics. That is, for each of these metrics applying noise reduction is beneficial in the majority of scenarios. These contrary results can likely be attributed to the positive effect of minimizing high-frequency fluctuations in noisy data sets on the derivation of these metrics; by reducing these fluctuations, these particular metrics can be more accurately derived.

Chapter Six: Conclusions and Recommendations

The remote sensing of vegetation phenology often relies on satellite-derived multi-temporal data sets of the normalized difference vegetation index (NDVI) in the form of per-pixel time series, for the characterization of surface vegetation health, growth and development. The effective use of such data sets, however, is hindered by the prevalence of noise that can be attributed principally to high-frequency variations in atmospheric conditions, such as cloud and aerosol presence, and from varying sun-sensor-surface angle geometries. Despite the notable quantity of research concerning this issue and the multitude of NDVI time series noise reduction strategies found in the literature, no unifying, standard method for dealing with such noise currently exists, nor is there a comprehensive review of the relevant literature. Existing comparisons between the application and performance of existing techniques are also limited, both in number and in scope.

The current research addresses the lack of both a comprehensive literature review and of comprehensive comparative evaluations of the existing NDVI time series noise reduction techniques. Guided by a series of informal research questions, this work was designed to address the following four objectives:

- Performing a comprehensive review of the literature concerning the acquisition and processing of NDVI time series, including noise reduction techniques, and the subsequent extraction of phenological variables. It was asked: what techniques are available? Were and when are they used? What is their success in the literature? What selection of techniques would be most appropriate for an empirical comparison in the current research?

- Determining which of the selected noise reduction techniques is most effectively applied to the present data set and study area. It was asked: is one technique superior to all other selected techniques? If so, which one? If not, why?

- Exploring factors that influence the effective application of the selected noise reduction techniques. It was asked: do land cover; year; the choice of

phenology-based NDVI metric; and level, strength and type of noise influence the performance of the selected techniques? Why or why not?

- Examining whether it is always best to apply NDVI time series noise reduction. It was asked: is noise reduction always beneficial? When is it, or is it not beneficial? Do land cover, noise level, year and choice of NDVI metric influence the beneficial nature of noise reduction?

Accomplishing these goals involved, first, the collection, organization and integration of the relevant literature concerning NDVI time series noise reduction into a comprehensive review, the results of which are presented in Chapter Two. From this review, six suitable candidate techniques were chosen for further research: the 4253H-Twice filter, the ARMD3-ARMA5 filter, the Asymmetric Gaussian and Double Logistic function-fitting techniques, the MVI filter and the Savitzky-Golay filter. An empirical comparison of these techniques comprised of root mean square error (RMSE) and time series metric-based evaluations were undertaken within the framework of a model environment. Six model time series were constructed from a multi-temporal data set of MODIS NDVI 16-day composites covering west-central Alberta, Canada, from 2003 through 2005, with noise introduced at random selections of dates in order to represent slightly (10%), moderately (40%) and very (70%) noisy time series.

Analysis of the results from the RMSE and metric evaluations revealed that, over the full range of conditions simulated in this experiment, both the Double Logistic and Asymmetric Gaussian function-fitting methods demonstrated the best overall ability to minimize noise in NDVI time series and preserve the integrity of the original NDVI signal; they portrayed a general superiority to the four candidate filtering techniques. The very similar capacities of the Double Logistic and Asymmetric Gaussian methods for approximating constant winter NDVI despite high frequency noise and their strict preservation of the upper envelope of NDVI values are likely reasons for their success. An examination of the stratified results revealed that while performance did vary with land cover, these patterns were most often caused by variations in the strength and type of introduced noise, rather than phenological differences between land cover types. In particular, the presence or absence of a negative bias in the noise led to varying levels of

over- or underestimation by techniques that presume a negative bias versus those that do not. The level of introduced noise also demonstrated notable influence on performance; some, like the 4253H-Twice filter, were better able to deal with minimal noise because they did not alter the remainder of the model time series as did other techniques. On the other hand, other techniques such as the Savitzky-Golay filter showed better performance with higher noise levels and were not as affected by the presence of considerable noise. Consideration of these effects is advisable when selecting a method for NDVI time series noise reduction; such selections should be guided by an examination of noise in the data set of interest, particularly with regard to strength or bias.

There appeared to be no meaningful annual variation in results, with similar patterns appearing through each of the three years evaluated. What little variation existed was largely explained by yearly changes in type and strength of noise. Notable, however, was the variable impact of noise reduction on individual phenological metrics. Indeed, the results showed such variability that few overall trends could be discerned. Nonetheless, more accurate derivations of key time-based metrics such as start of growing season, end of growing season, and length of growing season were generally observed from noise-reduced time series, while value-based metrics such as maximum NDVI and integrated NDVI were often better estimated by 'noisy' time series. It seems evident that noise reduction is an effective strategy for removing high-frequency fluctuations that might otherwise be mistaken for start- and end-of-season season events, but that these same procedures might also reduce, in an unpredictable fashion, our ability to extract some value-based metrics.

Investigating the overall benefit of noise reduction again produced quite varied and seemingly contradictory results. This contradiction arose from differences between the scenario-based and standardized performance score-based evaluation methods. In many cases, noise reduction was found to limit the subsequent ability to extract accurate phenological metrics, even though the overall accuracy of individual NDVI estimations might be improved. However, the probability that applying noise reduction would be beneficial was considerably influenced by one's choice of metric. The NDVI amplitude, average NDVI and maximum green-up metrics are the only metrics for which the

majority of scenarios demonstrated improved performance with noise reduction. Indeed, these were the only metric scenario results to show such patterns. This likely resulted from the positive effect of applying noise reduction and minimizing high frequency, negatively-biased fluctuations which could have caused inaccurate derivations of these particular metrics. This observation appeared contradictory to what was observed above regarding metrics, but can be attributed to a difference in evaluation techniques between the performance score-based assessment and the scenario analysis. It was concluded that while the derivation of particular value-based metrics was generally benefited by the application of noise reduction, where the extraction of key time-based metrics would be benefited by noise reduction the negative consequences of derivation from noisy time series were more serious than inaccuracies resulting from the potentially inhibiting application of noise reduction. As a result, users should be cautious about the use of noise reduction when the ultimate objective includes the extraction of phenological metrics.

By successfully addressing the four objectives and providing answers to the research questions posed above, the current work illustrates the difficulties and complexities of NDVI time series noise reduction. Those applying such data sets in the remote sensing of vegetation phenology are advised to carefully consider their objectives and particular characteristics of the data in question when addressing the issue of noise. The strength and presence of bias in the noise, and the decision to extract phenological variables from the data should be influential factors when selecting an approach to the reduction of noise in NDVI time series.

6.1 Achievements and Contributions

Not only did the current research successfully meet all four main objectives and provide answers to the research questions posed above, but through its successful completion, offers several important contributions to the current knowledge and understanding regarding NDVI time series noise reduction and the use of such time series in the remote sensing of vegetation phenology.

Compiling a review of the current literature concerning NDVI time series noise reduction provided not only a solid foundation for the subsequent empirical comparison

and selection of candidate techniques for further evaluation, but also, a much-needed integration of the relevant research into one comprehensive summary. The simple existence of such a summary creates the potential to advance current understandings of the topic by offering a basis for further comprehensive evaluations and comparisons of existing noise reduction strategies, not unlike what was successfully undertaken here. The mere importance of addressing the existence of noise in NDVI time series renders the present assembly of a literature review a valuable contribution. By providing a crucial opening for the possible establishment of standard, accepted approaches toward reducing such noise in NDVI time series, the integration of current knowledge and relevant literature is an important first step.

In addition, the empirical comparison of several candidate noise reduction techniques not only answered the research questions posed above, but also provided considerable insight into the factors influencing the application of these techniques and the numerous complexities involved. Overall results suggested that the two function-fitting techniques were generally superior to the four alternative filtering techniques for reducing noise in NDVI time series. However, the considerable influence that the level, strength and type of noise and one's choice of metric in particular, had on the successful application of all the selected noise reduction techniques clearly illustrated the need for a greater understanding of such strategies and their application. The circumstances in which noise reduction is applied can greatly influence its effectiveness, and the present project is one of very few investigations to address this concern, making it a vital addition to the current literature. In addition, the use of a controlled model environment, an approach that is rarely found, if at all, in the present literature, allowed for a more exhaustive exploration of the selected noise reduction techniques than is yet encountered. By knowing the original, model signal that exists prior to the introduction of noise, a more effective comparison of the noise-reduced and 'true' time series could be made than would be possible using real-world time series. The noise reduction of NDVI time series is evidently not simple and straightforward, and the present research is thus all the more valuable a contribution to current knowledge and understanding regarding the topic.

As well as providing further knowledge of the complexities of NDVI time series noise reduction, the present research also offers insight into the circumstances and situations in which the application of noise reduction techniques may *not* be beneficial. Indeed, questioning the advantage of applying such techniques is something that is not generally done in the literature, yet is undeniably important to their proper application. The current investigation is thus an essential source of further knowledge concerning this issue, particularly in light of the findings. It was revealed that the benefit of applying noise reduction is dependent upon one's objectives and subsequent use of NDVI time series. That is, if general noise reduction is desired - in other words, the minimization of the high-frequency fluctuations observed in noisy time series - the selected strategies are effective more often than not. However, when the subsequent extraction of NDVI time series metrics is the objective, the value of these techniques is less certain. Knowledge, even of the potential for noise reduction to lead to less desirable results, is crucial to the proper application and use of noise reduction strategies. In furthering this knowledge, the present research makes additional, fundamental contributions to current understandings of NDVI time series noise reduction.

6.2 Recommendations for Future Research

As with any respectable scientific investigation, the endeavour to answer a set of research questions led to the emergence of additional questions, which, though not answerable within the scope of the current research, may provide direction for avenues of future research. Primary among these additional questions is whether similar findings would result from a comparable examination of noise reduction strategies as applied to real-world NDVI time series. A similar investigation employing actual NDVI time serious would prove highly valuable as a means of verifying the results of the current work, particularly beyond the scope of the present model environment and study area. Were the costs and difficulties associated with collecting and employing the required ground truth data minimized, perhaps through the use of proxy data or through collaboration with larger institutions or organizations, such research might be more easily undertaken and would undoubtedly contribute additional knowledge and understanding regarding noise reduction in NDVI time series. For example, such an undertaking might

incorporate ground-based flux tower observations acquired through the international FLUXNET project as a surrogate for phenological ground truth.

Another recommendation that arises from the present research is that, if a similar empirical comparison of noise reduction strategies for NDVI time series were to be undertaken, a greater contribution could be made with the use of greater statistical rigour. As explained in Chapter Four, the results of the present analyses are very relative in nature. They are interval-level data from which no absolute measure of performance can be drawn, and to which no statistical significance testing could be easily applied. Although they are extremely informative, the nature of the results limit the current capacity for drawing conclusions regarding the six candidate noise reduction techniques and their effective application to multi-temporal NDVI data, to a certain extent. Incorporating absolute measures of performance and statistical significance testing into future work would be highly recommended. For example, analysis of the residuals and testing for the presence of autocorrelation would provide additional measures of noise reduction technique suitability and performance.

Another avenue of investigation that requires attention and is highlighted by the present research is continued work on the effects of NDVI time series noise reduction on subsequent derivation of vegetation phenology metrics. While the current study reveals the potential for inaccuracy as a result of applying such noise reduction techniques, more detailed analysis of the relationship between these techniques and the extraction of metrics is required before an adequate understanding of these effects can be reached. In particular, exploring this relationship using real-world NDVI time series and the necessary ground truth data would be especially pertinent to an understanding of this relationship and to the proper application of noise reduction strategies to NDVI data sets from which metrics are to be calculated. The current research represents a preliminary investigation into the effects of noise reduction on metric derivation, and perhaps provides a foundation on which further studies could be based, but such studies are needed if the implications of applying noise reduction for the subsequent calculation of NDVI time series metrics can be clearly understood.

A fourth and final recommendation can be made with regard to the future application of these and the results of similar investigations to current ecological, biological and climatological research objectives. The value of incorporating synoptic remotely-sensed measures of phenology derived from multi-temporal NDVI data sets into such objectives is by this time widely-recognized. However, superior knowledge and understanding concerning the proper application of noise reduction to such data sets is crucial to the effective use of noise reduction strategies and in the application of remotely-sensed vegetation phenology to these and other related research initiatives. While further investigation into the noise reduction of NDVI time series is important, the application of such knowledge to a range of multi-disciplinary scientific objectives is equally valuable. For instance, despite the key role that vegetation phenology plays in animal movement and habitat selection, these issues are rarely addressed by contemporary wildlife ecology research and management activities in a sophisticated manner. The current remote sensing-based efforts undertaken by the Foothills Model Forest Grizzly Bear Research Program (FMFGBRP) rely largely on static landscape attributes such as land cover and vegetation structure for the mapping and modeling of grizzly habitat. The incorporation of dynamic measures of vegetation phenology into these mapping and modeling activities would likely enhance the conservation efforts of the FMFGBRP.

The current research provides a crucial first step toward more a comprehensive understanding of the remote sensing of vegetation phenology, and more particularly, the application of noise reduction techniques to NDVI time series. It offers a framework for the subsequent pursuit of further knowledge, and a foundation for its future advancement.

References

ACHUFF, P. L., 1992, *Natural Regions, Subregions, and Natural History Themes of Alberta: A Classification for Protected Areas Management*. Alberta Environmental Protection, Edmonton, Alberta. 72 pp.

ACHUFF, P. L. and WALLIS, C., 1977, *A proposed policy for ecological reserves in Alberta*. Alberta Energy and Natural Resources, Edmonton, Alberta. 67 pp.

AHAS, R., AASA, A., MENZEL, A., FEDOTOVA, V. G. and SHEIFINGER, H., 2002, Changes in European spring phenology. *International Journal of Climatology*, **22**, 1727-1738.

AHL, D. E., GOWER, S. T., BURROWS, S. N., SHABANOV, N. V., MYNENI, R. B. and KNYAZIKHIN, Y., 2006, Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. *Remote Sensing of Environment*, **104**, 88-95.

ALHAMAD, M. N., STUTH, J. and VANNUCCI, M., 2007, Biophysical modelling and NDVI time series to project near-term forage supply: spectral analysis aided by wavelet denoising and ARIMA modelling. *International Journal of Remote Sensing*, **28**, 2513-2548.

ANDRES, L., SALAS, W. A. and SKOLE, D., 1994, Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. *International Journal of Remote Sensing*, **15**, 1115-1121.

AURELA, M., TOUVINEN, J.-P. and LAUILA, T., 2001, Net CO_2 exchange of a subarctic mountain birch system. *Theoretical and Applied Climatology*, **70**, 135-148.

AZZALI, S. and MENENTI, M., 2000, Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing*, **21**, 973-996.

BADECK, F.-W., BONDEAU, A., BÖTTCHER, K., DOKTOR, D., LUCHT, W., SCHABER, J. and SITCH, S., 2004, Responses of spring phenology to climate change. *New Phytologist*, **162**, 295 309.

BADHWAR, G. D., CARNES, J. G. and AUSTIN, W. W., 1982, Use of Landsat-derived temporal profiles from corn-soybean feature extraction and classification. *Remote Sensing of Environment*, **12**, 57-79.

BADHWAR, G. D., 1984, Use of LANDSAT-derived profile features for spring small-grains classification. *International Journal of Remote Sensing*, **5**, 783-797.

BECK, P. S. A., ATZBERGER, C., HØGDA, K. A., JOHANSEN, B. and SKIDMORE, A. K., 2006, Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODIS NDVI. *Remote Sensing of Environment*, **100**, 321-334.

BECK, P. S. A., JÖNSSON, P., HØGDA, K.-A., KARLSEN, S. R., EKLUNDH, L., and SKIDMORE, A. K., 2007, A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. *International Journal of Remote Sensing*, **28**, 4311-4330.

BLONDEL, J., DIAS, P.C., MAISTRE, M. and PERRET, P.., 1993, Habitat heterogeneity and life history variation of the Mediterranean blue tits (*Parus caeruleus*). *Auk*, **110**, 511-520.

BRADLEY, B. A., JACOB, R. W., HERMANCE, J. F. and MUSTARD, J. F., 2007, A curve fitting procedure to derive inter-annual phonologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, **106**, 137-145.

BRUCE, L. M., MATHUR, A. and BYRD, J. D. JR., 2006, Denoising and wavelet-based feature extraction of MODIS multi-temporal vegetation signatures. *GIScience and Remote Sensing*, **43**, 67-77.

CABRAL, A., DE VASCONCELOS, M. J. P., PEREIRA, J. M. C., BARTHOLOMÉ, E. and MAYAUX, P., 2003, Multi-temporal compositing approaches for SPOT-4 VEGETATION. *International Journal of Remote Sensing*, **24**, 3343-3350.

CARREIRAS, J. M. B., PEREIRA, J. M. C., SHIMABUKURO, Y. E. and STROPPIANA, D., 2003, Evaluation of compositing algorithms over the Brazilian Amazon using SPOT-4 VEGETATION data. *International Journal of Remote Sensing*, **24**, 3427-3440.

CHEN, J., JÖNSSON, P., TAMURA, M., GU, Z., MATSUSHITA, B. and EKLUNDH, L., 2004, A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, **91**, 332-344.

CHEN, P.-Y., SRINIVASAN, R., FEDOSEJEVS, G., and KINIRY, J. R., 2003, Evaluating different NDVI composite techniques using NOAA-14 AVHRR data. *International Journal of Remote Sensing*, **24**, 3403-3412.

CHEN, W. J., BLACK, T. A., YANG, P. C., BARR, A. G., NEUMANN, H. H., NESIC, Z., BLANKEN, P. D., NOVAK, M. D., ELEY, J., KETLER, J. and CUENCA, R., 1999, Effects of climate variability on the annual carbon sequestration by a boreal aspen forest. *Global Change Biology*, **5**, 41-53.

CHEN, X., TAN, Z., SCHWARTZ, M. D. and XU, C., 2000, Determining the growing season of land vegetation on the basis of plant phenology and satellite data in Northern China. *International Journal of Biometeorology*, **44**, 97-101.

CIHLAR, J., MANAK, D. and VOISIN, N., 1994, AVHRR bidirectional reflectance effects and compositing. *Remote Sensing of Environment*, **48**, 77-88.

CIHLAR, J., LY, H., LI, Z., CHEN, J., POKRANT, H. and HUANG, F., 1997, Multitemporal, Multichannel AVHRR Data Sets for Land Biosphere Studies – Artifacts and Corrections. *Remote Sensing of Environment*, **60**, 35-57.

CIHLAR, J., LATIFOVIC, R., CHEN, J., TRISHCHENKO, A., DU, Y., FEDOSEJEVS, G. and GUINDON, B., 2004, Systematic corrections of AVHRR image composites for temporal studies. *Remote Sensing of Environment*, **89**, 217-233.

COWAN, F. L. and ODELL, G. H., 1990, More on estimating tillage effects: reply to Dunnell and Yorston. *American Antiquity*, **55**, 598-605.

DEERING, D. W., 1978, Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. Ph.D. Dissertation, Texas A & M University, College Station, TX, 338 pp.

DEFRIES, R. S. and TOWNSHEND, 1994, NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, **15**, 3567-3586.

DEFRIES, R., HANSEN, M. and TOWNSHEND, J., 1995, Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment*, **54**, 209-222.

DELBART, N., KERGOAT, L., LE TOAN, T., LHERMITTE, J. and PICARD, G., 2005, Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment*, **97**, 26-38.

DELBART, N., LE TOAN, L., KERGOAT, L. and FEDOTOVA, V., 2006, Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT_VGT data (1982-2004). *Remote Sensing of Environment*, **101**, 52-62.

DESPLAND, E., ROSENBERG, J. and SIMPSON, S. J., 2004, Landscape structure and locust swarming: a satellite's eye view. *Ecography*, **27**, 381-391.

DOUGILL, A. and TRODD, N., 1999, Monitoring and modelling open savannas using multisource information: analyses of Kalahari studies. *Global Ecology and Biogeography*, **8**, 211-221.

FANG, H. –T. and HUANG, D. –S., 2004, Noise reduction in lidar signal based on discrete wavelet transform. *Optics Communications*, **233**, 67-76.

FENSHOLT, R., SANDHOLT, I., and RASMUSSNEN, M. S., 2004, Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment*, **91**, 490-507.

FILIPOVA-RACHEVA, D. and HALL-BEYER, M., Smoothing of NDVI time series curves for monitoring of vegetation changes in time. University of Calgary, Unpublished. 4 pp.

FISCHER, A., 1994a, A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Environment*, **48**, 220-230.

FISCHER, A., 1994b, A simple model for the temporal variations of NDVI at regional scale over agricultural countries. Validation with ground radiometric measurements. *International Journal of Remote Sensing*, **15**, 1421-1446.

FISHER, J. I., MUSTARD, J. F. and VADEBONCOEUR, M. A., 2006, Green leaf phenology at Landsat resolution: scaling from the field to satellite. *Remote Sensing of Environment*, **100**, 265-279.

GAGE, S. H., ISARD, S. A. and COLUNGA-G., M., 1999, Ecological scaling of aerobiological dispersal processes. *Agricultural and Forest Meteorology*, **97**, 249-261.

GAILLARD, J.-M., FESTA-BIANCHET, M., YOCCOZ, N. G., LOISON, A. and TOIGO, C., 2000, Temporal variation in fitness components and population dynamics of large herbivores. *Annual Review of Ecology and Systematics*, **31**, 367-393.

GAO, Z., HUETE, A. R. and DIDAN, K., 2003, Multisensor comparisons and validation of MODIS vegetation indices at the semiarid Jordana experimental range. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 2368-2381.

GILABERT, M. A., GONZÁLEZ-PIQHERAS, J., GARCÍA-HARO, F. J. and MELIÁ, J., 2002, A generalized soil-adjusted vegetation index. *Remote Sensing of Environment*, **82**, 303-310.

GOVERNMENT OF ALBERTA, 2005 Natural regions and subregions of Alberta. Available online at: http://tprc.alberta.ca/parks/heritageinfocentre/naturalregions/default.aspx (accessed 12 August 2007).

GOWARD, S. N., DYE, D. G. and TUCKER, C. J., 1985, North American vegetation patterns observed by NOAA-7 AVHRR, *Vegetatio*, **64**, 3.

GOWARD, S. N., MARKHAM, B. and DYE, D., 1991, Normalized Difference Vegetation Index measurements from the Advanced Very High Resolution Radiometer. *Remote Sensing of Environment*, **35**, 257-277.

GUTMAN, G. G., 1991, Vegetation indices from AVHRR: an update and future prospects. *Remote Sensing of Environment*, **35**, 121-136.

GUPTA, R. K., PRASAD, S., SESHA SAI, M. V. R. and VISWANADHAM, T. S., 1997, The estimation of surface temperature over an agricultural area in the state of Haryana and Panjab, India, and its relationship with the Normalized Difference Vegetation Index (NDVI), using NOAA-AVHRR data. *International Journal of Remote Sensing*, **18**, 3729-3741.

HALL, C. A. S., EKDAHL, C. A. and WARTENBERG, D. E., 1975, A fifteen-year record of biotic metabolism in the Northern Hemisphere. *Nature*, **255**, 136-138.

HERMANCE, J. F., 2007, Stabilizing high-order, non-classical harmonic analysis of NDVI data for average annual models by damping model roughness. *International Journal of Remote Sensing*, **28**, 2801-2819.

HERMANCE, J.F., JACOB, R. W., BRADLEY, B. A. and MUSTARD, J. F., 2007, Extracting phenological signals from multi-year AVHRR NDVI time series: framework for applying high-order annual splines with roughness damping. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 3264-3276.

HOBSON, D., 2005, Denning of Grizzly Bears in the Foothills Model Forest. In *Foothills Model Forest Grizzly Bear Research Program 1999-2003 Final Report*, G. B. Stenhouse and K. Graham (Eds.), pp. 32-37.

HOLBEN, B. N., and FRASER, R. S., 1984, Red and near-infrared response to off-nadir viewing. *International Journal of Remote Sensing*, **5**, 145-160.

HOLBEN, B. N., 1986, Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, **7**, 1417-1434.

HUETE, A. R., 1988, A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, **25**, 295-309.

HUETE, A. R. and LIU, H. Q., 1994, An error and sensitivity analysis of the atmospheric- and soil correcting variants of the NDVI for MODIS-EOS. *IEEE Transactions on Geoscience and Remote Sensing*, **32**, 897-905.

HUETE, A. R., LIU, H. Q., BATCHILY, K. and VAN LEEUWEN, W., 1997, A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, **59**, 440-451.

HUETE, A., JUSTICE, C. and VAN LEEUWEN, W., 1999. *MODIS Vegetation Index (MOD 13) Algorithm Theoretical Basis Document*. Version 3. Available online at: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf (accessed 23 March 2006).

HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E.P., GAO, X. and FERREIRA, L.G., 2002, Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, **83**, 195-213.

HURLBERT, A. H. and HASKELL, J. P., 2003, The effect of energy and seasonality on avian species richness and community composition. *The American Naturalist*, **16**, 83-97.

JAKUBAUSKAS, M. E., LEGATES, D. R. and KASTENS, J. H., 2001, Harmonic analysis of time series AVHRR NDVI data. *Photogrammetric Engineering and Remote Sensing*, **67**, 461 470.

JAKUBAUSKAS, M. E., PETERSON, D. L., KASTENS, J. H. and LEGATES, D. R. 2002, Time series remote sensing of landscape-vegetation interactions in the southern Great Plains. *Photogrammetric Engineering and Remote Sensing*, **68**, 1021-1030.

JENSEN, J. R., 1996, Image enhancement. In *Introductory Digital Image Processing: A Remote Sensing Perspective*, edited by Clarke, K. C. (Upper Saddle River, New Jersey: Prentice Hall), pp.139-196.

JENSEN, J. R., 2000, Remote sensing of vegetation. In *Remote Sensing of the Environment: An Earth Resource Perspective*, edited by K. C. Clarke (Upper Saddle River, New Jersey: Prentice Hall), pp. 333-377.

JIA, G. J., EPSTEIN, H. E. and WALKER, D. A., 2004, Controls over intra-seasonal dynamics of AVHRR NDVI for the Arctic tundra in northern Alaska. *International Journal of Remote Sensing*, **25**, 1547-1564.

JÖNSSON, P. K. and EKLUNDH, L., 2002, Seasonality Extraction by Function-fitting to Time Series of Satellite Sensor Data. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 1824-1832.

JÖNSSON, P. and EKLUNDH, L., 2004, TIMESAT – a program for analyzing time-series of satellite sensor data. *Computers and Geosciences*, **30**, 833-845.

JÖNSSON, P. and EKLUNDH, L., 2006, TIMESAT – a program for analyzing time-series of satellite sensor data: Users Guide for TIMESAT 2.3. (Sweden: Malmö and Lund), 39 pp.

JORDAN, C. F., 1969, Derivation of leaf area index from quality of light on the forest floor, *Ecology*, **50**, 663-666.

JUSTICE, C. O., TOWNSHEND, J. R. G., HOBLEN, B. N. and TUCKER, C. J., 1985, Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal f Remote Sensing*, **6**, 1271-1318.

JUSTICE, C. O., HOLBEN, B. N. and GWYNNE, M. D., 1986, Monitoring East African vegetation using AVHRR data. *International Journal of Remote Sensing*, **7**, 1453-1474.

JUSTICE, C. O., TOWNSHEND, J. R. G., VERMOTE, E. F., MASUOKA, E., WOLFE, R. E., SALEOUS, N., ROY, D. P. and MORISETTE, J. T., 2002, An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, **83**, 3-15.

KANG, S., RUNNING, S. W., LIM, J. -H., ZHAO, M., PARK, C. -R. and LOEHMAN, R., 2003, 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**, 232-242.

KANG, S., RUNNING, S. W., ZHAO, M., KIMBALL, J. S. and GLASSY, J., 2005, Improving continuity of MODIS terrestrial photosynthesis products using an interpolation scheme for cloudy pixels. *International Journal of Remote Sensing*, **26**, 1659-676.

KATHUROJU, N., WHITE, M., SYMANZIK, J., SCHWARTZ, M. D., POWELL, J. A. and NEMANI, R. R., 2007, On the use of the advance very high resolution radiometer for development of prognostic land surface phenology models. *Ecological Modelling*, **201**, 144-156.

KNIGHT, J. F., LUNETTA, R. S., EDIRIWICKREMA, J. and KHORRAM, S., 2006, Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: a phenology- based approach. *GIScience & Remote Sensing*, **43**, 1-23.

KOBAYASHI, H. and DYE, D. G., 2005, Atmospheric conditions for monitoring the long-term dynamics in the Amazon using normalized difference vegetation index. *Remote Sensing of Environment*, **97**, 519-525.

LAND PROCESSES DATA ACTIVE ARCHIVE CENTRE (LP DAAC), 2005, MODIS Data Products. Available online at: http://lpdaac.usgs.gov/modis/mod13q1v4.asp (accessed 12 April 2006).

LEE, R., YU, F., PRICE, K. P., ELLIS, J. and SHI, P., 2002, Evaluating vegetation phenological patterns in Inner Mongolia using NDVI time-series analysis. *International Journal of Remote Sensing*, **23**, 2505-2512.

LI, Z. and STRAHLER, A. H., 1992, Geometric-optical bidirectonal reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 276-292.

LILLESAND, T. M., KIEFER, R. W. and CHIPMAN, J. W., 2008, *Remote Sensing and Image Interpretation*, Hoboken, New Jersey: John Wiley & Sons.

LLOYD, D., 1990, A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, **11**, 2269-2279.

LOE, L. E., BONENFANT, C., MYSTERUD, A., GAILLARD, J.-M., LANGVATN, R., KLEIN, F., CALENGE, C., ERAGON, T., PETTORELLI, N., and STENSETH, N. C., 2005, Climate predictability and breeding phenology in red deer: timing and synchrony of rutting and calving in Norway and France. *Journal of Animal Ecology*, **75**, 579-588.

LOVELL, J. L. and GRAETZ, R. D., 2001, Filtering pathfinder AVHRR land NDVI data for Australia. *International Journal of Remote Sensing*, **22**, 2649-2654.

LU, X., LIU, R., LIU J. and LIANG, S., 2007, Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products. *Photogrammetric Engineering and Remote Sensing*, **73**, 1129-1139.

MA, M. and VEROUSTRAETE, F., 2006, Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China. *Advances in Space Research*, **37**, 835-840.

MACDONALD, R. B. and HALL, F. G., 1980, Global crop forecasting. *Science*, **208**, 670-679.

MAXWELL, S. K., HOFFER, R. M. and CHAPMAN, P. L., 2002, AVHRR composite period selection for land cover classification. *International Journal of Remote Sensing*, **23**, 5043-5059.

MCDERMID, G., PAPE, A. and LASKIN, D., 2006, Remote sensing: map production update. In *Foothills Model Forest Grizzly Bear Research Program 2005 Annual Report*, edited by G. Stenhouse and K. Graham (Hinton, Alberta, Canada), pp. 28-33.

MCGREW, J. C. JR., and MONROE, C. B., 2000, *An Introduction to Statistical Problem Solving in Geography*, 2nd Ed. (Boston, MA: McGraw-Hill).

MENENTI, M., AZZALI, S., VERHOEF, W. and VAN SWOL, R., 1993, Mapping agroecological zones and time lag in vegetation growth by means of Fourier

analysis of time series of NDVI images. *Advances in Space Research*, **13**, (5)233-(5)27.

MENZEL, A., ESTRELLA, N. and FABIAN, P., 2001, Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. *Global Change Biology*, **7**, 657-666.

MIURA, T., HUETE, A. R. and YOSHIOKA, H., 2000, Evaluation of sensor calibration uncertainties on vegetation indices for MODIS. *IEEE Transactions on Geoscience and RemoteSensing*, **38**, 1399-1409.

MOODY, A. and JOHNSON, D. M., 2001, Land-surface phonologies from AVHRR using the discrete fourier transform. *Remote Sensing of Environment*, **75**, 305-323.

MORISETTE, J. T., PRIVETTE, J. L. and JUSTICE, C. O., 2002, A framework for the validation of MODIS Land products. *Remote Sensing of Environment*, **83**, 77-96.

MOULIN, S., KERGOAT, L., VIOVY, N. and DEDIEU, G., 1997, Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. *Journal of Climate*, **10**, 1154-1170.

MUNRO, R. H. M., PRICE, M. H. H. and STENHOUSE, G. B., 2005, The Diet of Grizzly Bears, *Ursus arctos*, in West-Central Alberta, Canada. In *Foothills Model Forest Grizzly Bear Research Program 1999-2003 Final Report*, G. B. Stenhouse and K. Graham (Eds.), pp. 11-20.

MYNENI, R. B., HALL, F. G., SELLERS, P. J. and MARSHAK, A. L., 1995, The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 481-486.

MYNENI, R. B., KEELING, C. D., TUCKER, C. J., ASRAR, G. and NEMANI, R. R., 1997, Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, **386**, 698-702.

MYNENI, R. B., TUCKER, C. J., ASRAR, G. and KEELING, C. D., 1998, Interannual variations in satellite-sensed vegetation index data from 1981 to 1991. *Journal of Geophysical Research*, **103(D6)**, 6145-6160.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, 2006, Terra mission at five-year milestone in Earth check-up. *Terra: The EOS Flagship*. Available online at: http://terra.nasa.gov/ (accessed 10 April 2006).

NORWINE, J. and GREEGOR, D. H., 1983, Vegetation classification based on advanced very high resolution radiometer (AVHRR) satellite imagery. *Remote Sensing of Environment*, **13**, 69-87.

NUMATA, I., ROBERTS, D. A., CHADWICK, O. A., SCHIMEL, J., SAMPAIO, F. R., LEONIDAS, F. C. and SOARES, J. V., 2007, Characterization of pasture biophysical properties and the impact of grazing intensity using remotely sensed data. *Remote Sensing of Environment*, **109**, 314-327.

OLSSON, L. and EKLUNH, L., 1994, Fourier series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing*, **15**, 3735-3741.

PAÑUELAS, J., FILELLA, I., ZHANG, X., LLORENS, OGAYA, R., LLORET, F., COMAS, P., ESTIARTE, M. and TERRADAS, J., 2004, Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist*, **161**, 837-846.

PARKINSON, C. L., 2003, Aqua: An Earth-Observing Satellite Mission to Examine Water and Other Climate Variables. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 173-183.

PARUELO, J. M. and LAUENROTH, W. K., 1998, Interannual variability of NDVI and its relationship to climate for North American shrublands and grasslands. *Journal of Biogeography*, **25**, 721-733.

PENFOUND, W. T., HALL, T. F. K and HESS, D., 1945, The spring phenology of plants in and around reservoirs in north Alabama with particular reference to malaria control. *Ecology*, **26**, 332-352.

PETTORELLI, N., VIK, J. O., MYSTRERUD, A., CAILLARD, J. M., TUCKER, C.J. and STENSETH, N. C., 2005, Using the satellite-derived NDVI to assess ecological responses to environmental change. *TRENDS in Ecology and Evolution*, **20**, 503-510.

PIAO, S., FANG, J., ZHOU, L., CIAIS, P. and ZHU, B., 2006, Variations in satellite-derived phenology of China's temperate vegetation. *Global Change Biology*, **12**, 672-685.

PRINCE, S. D., 1991, Satellite remote sensing of primary production; comparison of results for Sahelian grasslands 1981-1988. *International Journal of Remote Sensing*, **12**, 1301-1311.

RABINER, L. R., SAMBUR, M. R. and SCHMIDT, C. E., 1975, Applications of a nonlinear smoothing algorithm to speech processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **23**, 552-557.

RAHMAN, H. and DEDIEU, G., 1994, SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, **15**, 123-143.

REED, B. C., 2006, Trend analysis of time-series phenology of North America derived from satellite data. *GIScience and Remote Sensing*, **43**, 24-38.

REED, B. C., BROWN, J. F., VANDERZEE, D., LOVELAND, T. R., MEREHANT, J. W. and OHLEN, D. O., 1994, Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, **5**, 703-714.

REED, B. C., WHITE, M. and BROWN, J. F., 2003, Remote Sensing Phenology. In *Phenology: An Integrative Environmental Science*, edited by Mark D. Schwartz (Dordrecht, the Netherlands: Kluwer Academic Publishers), pp. 365-381.

REEVES, M. C., ZHAO, M. and RUNNING, S. W., 2006, Applying improved estimates of MODIS productivity to characterize grassland vegetation dynamics. *Rangeland Ecology & Management*, **59**, 1-10.

RICOTTA, C., REED, B. C. and TIESZEN, L. T., 2003, The role of C_3 and C_4 grasses to interannual variability in remotely sensed ecosystem performance over the US Great Plains. *International Journal of Remote Sensing*, **24**, 4421-4431.

ROERINK, G. J., MENENTI, M., and VERHOEF, W., 2000, Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, **21**, 1911-1917.

ROY, D. P., BORAK, J. S., DEVADIGA, S., WOLFE, R. E., ZHENG, M. and DESCLOITRE, J., 2002, The MODIS Land product quality assessment approach. *Remote Sensing of Environment*, **83**, 62-76.

RUNNING, S. W., JUSTICE, C. O., SALOMONSON, B., HALL, D., BARKER, J., KAUFMANN,Y. J., STRAHLER, A. H., HUETE, A. R., MULLER, J. –P., VENDERBILT, V., WAN, Z. M., TEILLET, P., and CARNEGGIE, D., 1994, Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing*, **15**, 3587-3620.

SAKAMOTO, T., YOKOZAWA, M., TORITANI, H., SHIBAYAMA, M., ISHITSUKA, N. and OHNO, H., 2005, A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, **96**, 366-374.

SAKAMOTO, T., VAN NGUYEN, N., OHNO, H., ISHITSUKA, N. and YOKOZAWA, M., 2006, Spatio temporal distribution of rice phenology and cropping systems in the Mekong Delta with special reference to the seasonal water flow of the Mekong and Bassac rivers. *Remote Sensing of Environment*, **100**, 1-16.

SANZ, J. J., POTTI, J., MORENO, J., MERINO, S. and FRÍAS, O., 2003, Climate change and fitness components of a migratory bird breeding in the Mediterranean region. *Global Change Biology*, **9**, 461-472.

SAVITZKY, A. and GOLAY, M. J. E., 1964, Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, 1627-1639.

SCHMIDT, K. S. and SKIDMORE, A. K., 2003, Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, **85**, 92-108.

SCHMIDT, K. S. and SKIDMORE, A. K., 2004, Smoothing vegetation spectra with wavelets. *International Journal of Remote Sensing*, **25**, 1167-1184.

SCHWARTZ, M. D., 1998, Green-wave phenology. *Nature*, **394**, 839-840.

SCHWARTZ, M. D. and REED, B. C., 1999, Surface phenology and satellite sensor-derived onset of greenness: and initial comparison. *International Journal of Remote Sensing*, **20**, 3451 3457.

SCHWARTZ, M. D. and REITER, B. E., 2000, Changes in North American spring. *International Journal of Climatology*, **20**, 929-932.

SCHWARTZ, M. D., REED, B. C. and WHITE, M. A., 2002, Assessing satellite-derived start-of season measures in the conterminous USA. *International Journal of Climatology*, **22**, 1793-1805.

SELLERS, P. J., TUCKER, C. J., COLLATZ, G. J., LOS, S. O., JUSTICE, C. O., DAZLICH, D. A. and RANDALL, D. A., 1994, A global 10 by 10 NDVI data set for global studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *International Journal of Remote Sensing*, **15**, 3519-3545.

SWETS, D. L., REED, B.C., ROWLAND, J. R. and MARKO, S. E., 1999, A weighted least-squares approach to temporal smoothing of NDVI. In *ASPRS Annual Conference, From Image to Information*, Portland, Oregon, May 17-21 1999, Bathesda, Maryland, (American Society for Photogrammetry and Remote Sensing).

SUZUKI, R., NOMAKI, T. and YASUNARI, T., 2003, West-east contrast of phenology and climate in northern Asia revealed using a remotely sensed vegetation index. *International Journal of Biometeorology*, **47**, 126-138.

TADDEI, R., 1997, Maximum Value Interpolated (MVI): a Maximum Value Composite method improvement in vegetation index profiles analysis. *International Journal of Remote Sensing*, **18**, 2365-2370.

TOWNSHEND, J. R. G. and JUSTICE, C. O., 1986, Analysis of the dynamics of African vegetation using the normalized difference vegetation index. *International Journal of Remote Sensing*, **7**, 1435-1445.

TUCKER, C. J., TOWNSHEND, J. R. G. and THOMAS, E. G., 1985, African land-cover classification using satellite data. *Science*, **227**, 369-375.

TUCKER, C. J., SLAYBACK, D. A., PINSON, J. E., LOS, S. O., MYNENI, R. B. and TAYLOR, M. G., 2001, Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999.

TUKEY, J. W., 1977, Robust nonlinear data smoothers; theory definitions and applications. PhD thesis, Princeton University, Princeton, NJ.

TURNER, D. P., RITTS, W. D., COHEN, W. B., GOWER, S. T., ZHAO, M., RUNNING, S. W., WOFSY, S. C., URBANSKI, S., DUNN, A. L. and MUNGER, J. W., 2003, Scaling Gross Primary Production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation. *Remote Sensing of Environment*, **88**, 256-270.

TURNER, D. P., RITTS, W. D., COHEN, W. B., GOWER, S. T., RUNNING, S. W., ZHAO, M., COSTA, M. H., KIRSCHBAUM, A. A., HAM, J. M., SALESKA, S. R. and AHL, D. E., 2006, Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment*, **102**, 282-292.

VAN DIJK, A., CALLIS, S. L., SAKAMOTO, C. M. and DECKER, W. L., 1987, Smoothing vegetation index profiles: an alternative method for reducing radiometric disturbance in NOAA/AVHRR data. *Photogrammetric Engineering and Remote Sensing*, **53**, 1059-1067.

VAN LEEUWEN, W. J. D., HUETE, A. R., DIDAN, K. AND LAING, T., 1997, Modeling bi-directional reflectance factors for different land cover types and surface components to standardize vegetation indices. In *Proceedings of the 7th International Symposium on Physical Measurements and Signatures in Remote Sensing*, Courcheval, France; 7-11 April, (G. Guyot and T. Phulpin, Eds.), A.A. Balkema, Roggerdam, pp. 373-380.

VAN LEEUWEN, W., HUETE, A. R. and LAING, T. W., 1999, MODIS vegetation index compositing approach: a prototype with AVHRR data. *Remote Sensing of Environment*, **69**, 264-280.

VELLEMAN, P. F., 1980, Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association*, **75**, 609-615.

VERBESSELT, J., LHERMITTE, S., COPPIN, P., EKLUNDH, L. and JÖNSSON, P., 2004, Biophysical drought metrics extraction by time series analysis of SPOT Vegetation data. In *IEEE International Geoscience and Remote Sensing Symposium, Science for Scoeity: Exploring and Managing a Changing Planet*,

Anchorage, Alaska, September 20-24 2004, (IEEE Geoscience and Remote Sensing Society).

VERHOEF, W., MENENTI, M. and AZZALI, S., 1996, A colour composite of NOAA-AVHRR NDVI based on time series analysis (1981-1992). *International Journal of Remote Sensing*, **17**, 231-235.

VERMOTE, E. F., EL SALEOUS, N. Z. and JUSTICE, C.O., 2002, Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment*, **83**, 97-111.

VICENTE-SERRANO, S. M., GRIPPA, M., DELBART, N., LE TOAN, T. and KERGOAT, L., 2006, Influence of seasonal pressure patterns on temporal variability of vegetation activity in central Siberia. *International Journal of Climatology*, **26**, 303-321.

VIOVY, N., ARINO, O. and BELWARD, A.S., 1992, The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, **13**, 1585-1590.

WAGENSEIL, H. and SAMIMI, C., 2006, Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. *International Journal of Remote Sensing*, **27**, 3455-3471.

WALTHALL, C.L., NORMAN, J. M., WELLES, J. M., CAMPBELL, G. and BLAD, B. L., 1985, Simple equation to approximate the bidirectional reflectance from vegetative canopies and bare soil surfaces. *Applied Optics*, **24**, 383-387.

WANG, Q., TENHUNEN, J., DINH, N. Q., REICHSTEIN, M., VESALA, T. and KERONEN, P., 2004, Similarities in ground- and satellite-based NDVI time series and their relationship to phenological activity of a Scots pine forest in Finland. *Remote Sensing of Environment*, **93**, 225-237.

WANG, Q., TENHUNEN, J., DINH, N. Q., REICHSTEIN, M., OTIENO, D., GRANIER, A. and PILEGARRD, K., 2005a, Evaluation of seasonal variation of MODIS derived leaf area index at two European deciduous broadleaf forest sites. *Remote Sensing of Environment*, **96**, 475-484.

WANG, Q., ADIKU, S., TENHUNEN, J. and GRANIER, A., 2005b, On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sensing of Environment*, **94**, 244-255.

WIEGAND, C. L., RICHARDSON, A. J. and KANEMASU, E. T., 1979, Leaf area index estimates for wheat from Landsat and their implications for evapotranspiration and crop modeling. *Agronomy Journal*, **71**, pp. 336.

WILLMOTT, C. J., 1982, Some comments on the evaluation of model performance. *Bulletin American Meteorological Society*, **63**, 1309-1313.

WILSON, K. B. and BALDOCCHI, D. D., 2000, Seasonal and interannual variability of energy fluxes over a broadleaved temperate deciduous forest in North America. *Agricultural and Forest Meteorology*, **100**, 1-18.

WHITE, M. A., THORNTON, P. E. and RUNNING, S. W., 1997, A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, **11**, 217-234.

WHITE, M. A., RUNNING, S. W. and THORNTON, P. E., 1999, The impact of growing season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern U.S. deciduous forest. *International Journal of Beiometeorology*, **42**, 139-145.

WHITE, M. A., NEMANI, R. R., THORNTON, P. E. and RUNNING, S. W., 2002, Satellite evidence of phenological differences between urbanized and rural areas of the eastern United States deciduous broadleaf forest. *Ecosystems*, **5**, 260-277.

XIAO, X., BOLES, S., LIU, J., ZHUANG, D. and LIU, M., 2002, Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. *Remote Sensing of Environment*, **82**, 335-348.

XIAO, X., HAGEN, S., ZHANG, Q., KELLER, M. and MOORE, B. III, 2006, Detecting leaf phenology of seasonally moist tropical forests in South America with multi-temporal MODIS images. *Remote Sensing of Environment*, **103**, 465-473.

YANG, W., YANG, L. and MERCHANT, J.W., 1997, An assessment of AVHRR/NDVI ecoclimatological relations in Nebraska, U.S.A. *International Journal of Remote Sensing*, **18**, 2161-2180.

YU, F., PRICE, K. P., ELLIS, J. and SHI, P., 2003, Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment*, **87**, 42-54.

ZHANG, X., FRIEDL, M. A., SCHAAF, C.B., STRAHLER, A. H., HODGES, J. C. F., GAO, F., REED, B. C. and HUETE, A., 2003, Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, **84**, 471-475.

ZHANG, X., FRIEDL, M. A., SCHAFF, C. B. and STRAHLER, A. H., 2004, Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. *Global Change Biology*, **10**, 1133-1145.

ZHOU, L., TUCKER, C. J., KAUFMANN, R. K., SLAYBACK, D., SHABANOV, N. V. and MYNENI, R. B., 2001, Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research*, **106(D17)**, 20,069-20,083.

APPENDIX A: ORIGINAL MODEL NDVI TIME SERIES AND THE INTRODUCTION OF 10%, 40% AND 70% NOISE

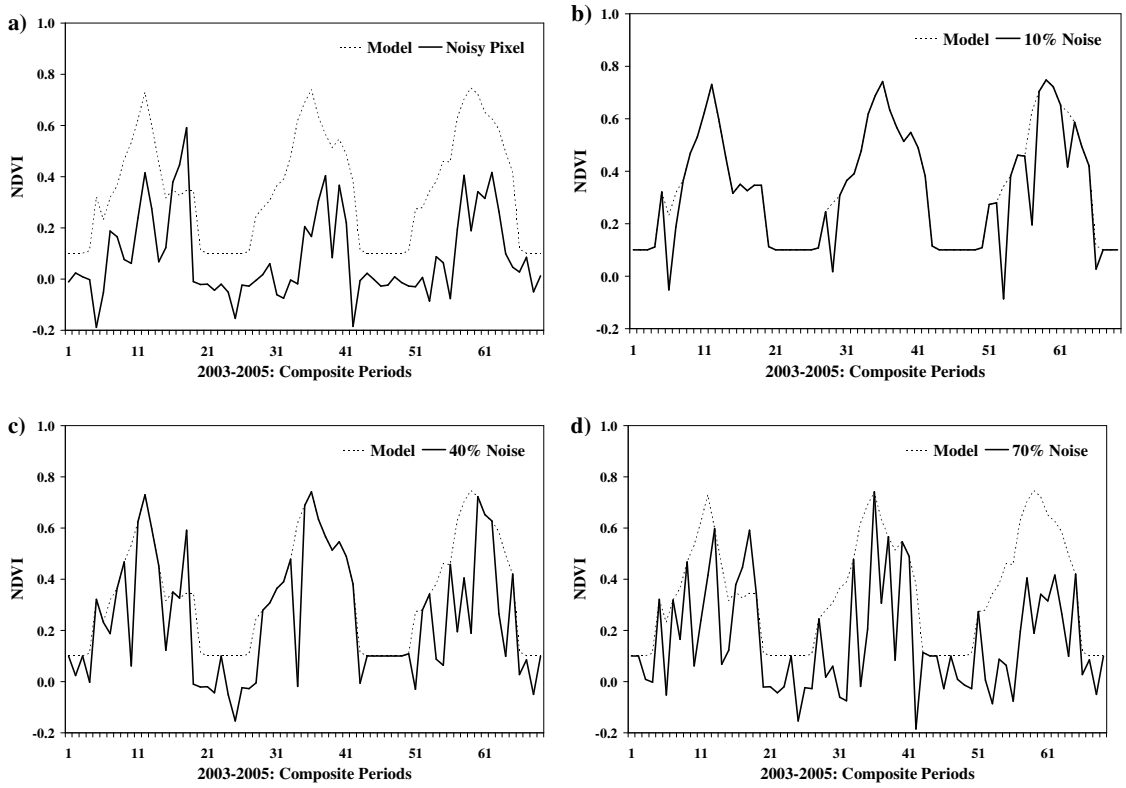

Figure A - 1: Grassland model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

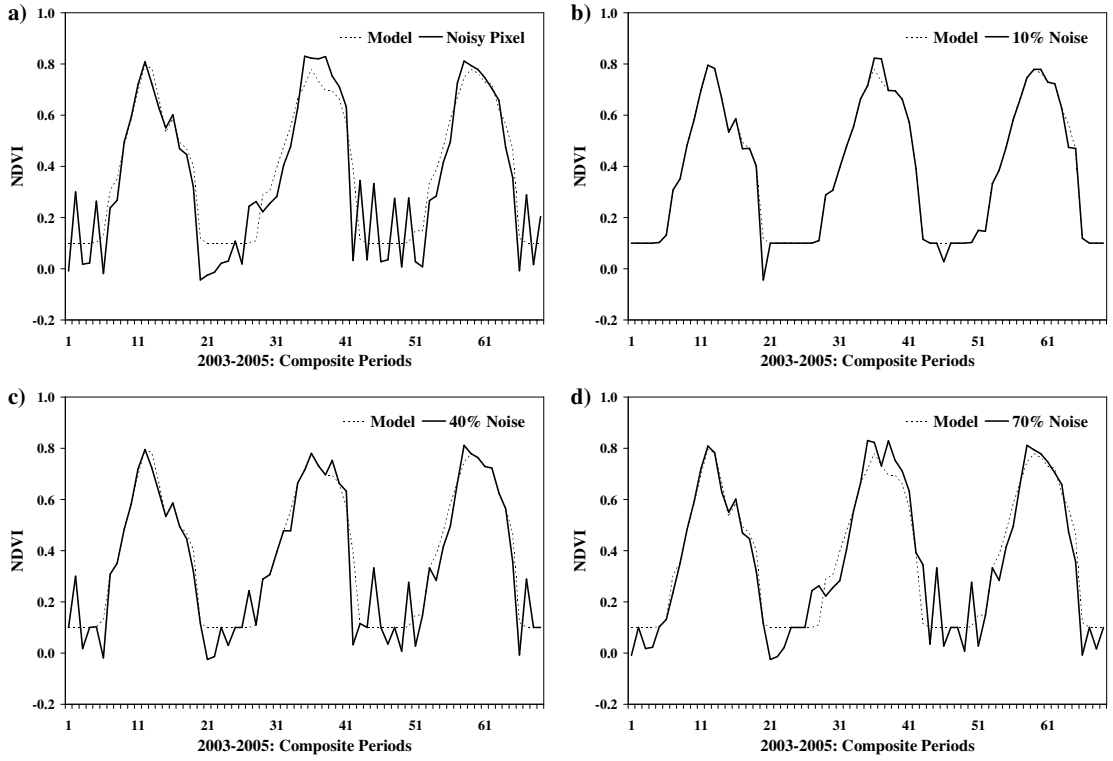

Figure A - 2: Parkland model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

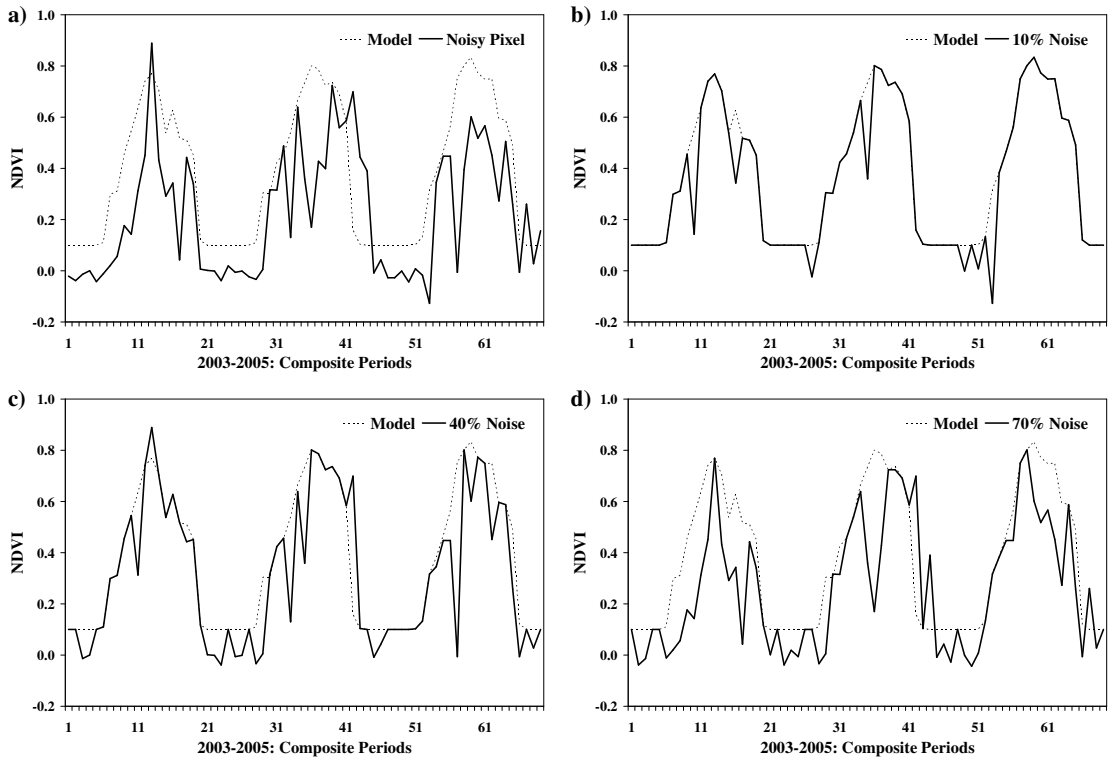

Figure A - 3: Boreal model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

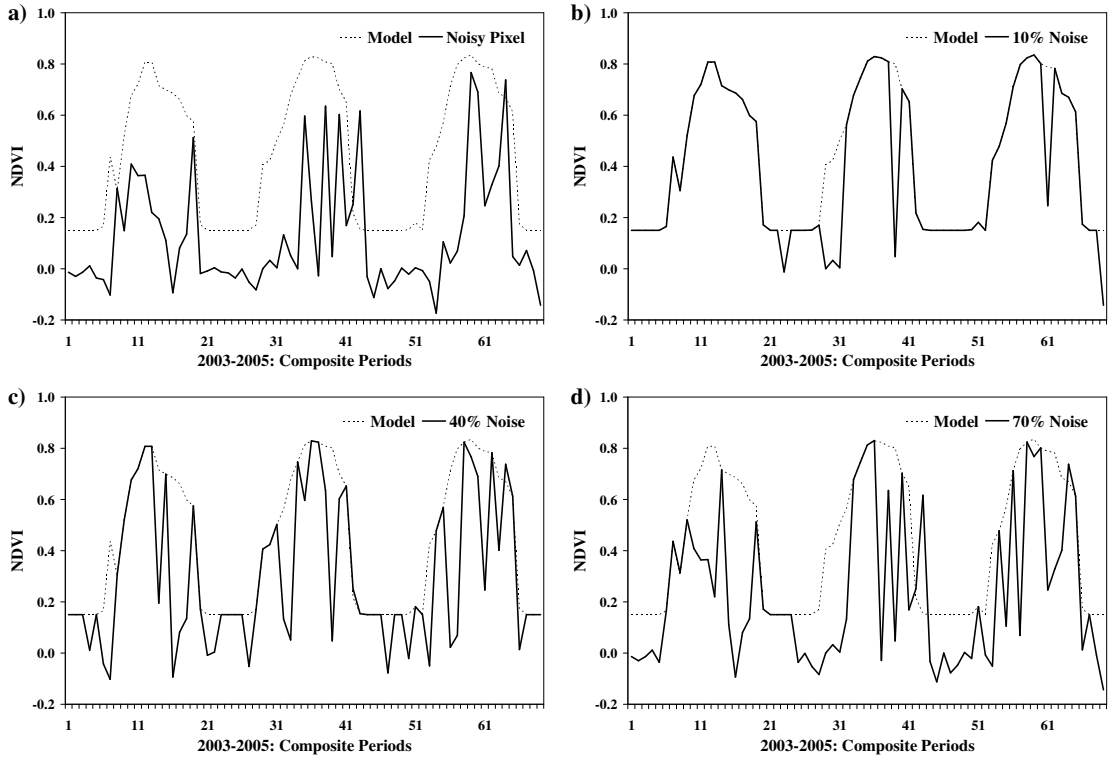

Figure A - 4: Lower Foothills model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

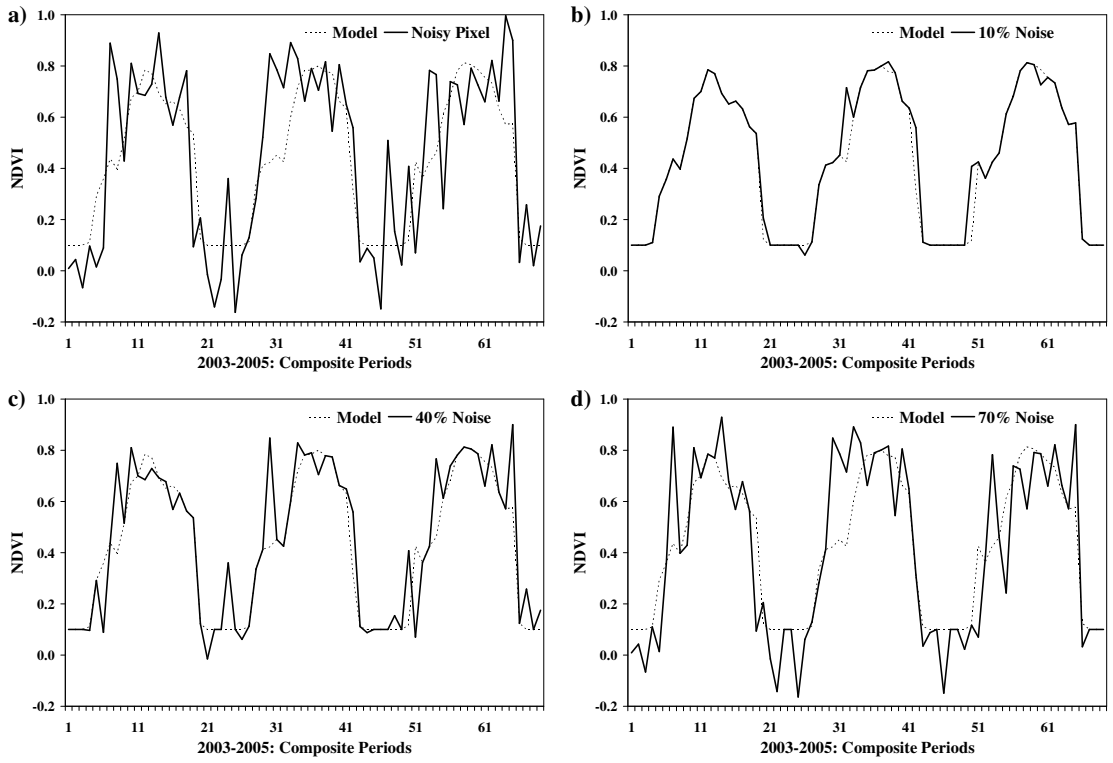

Figure A - 5: Montane model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

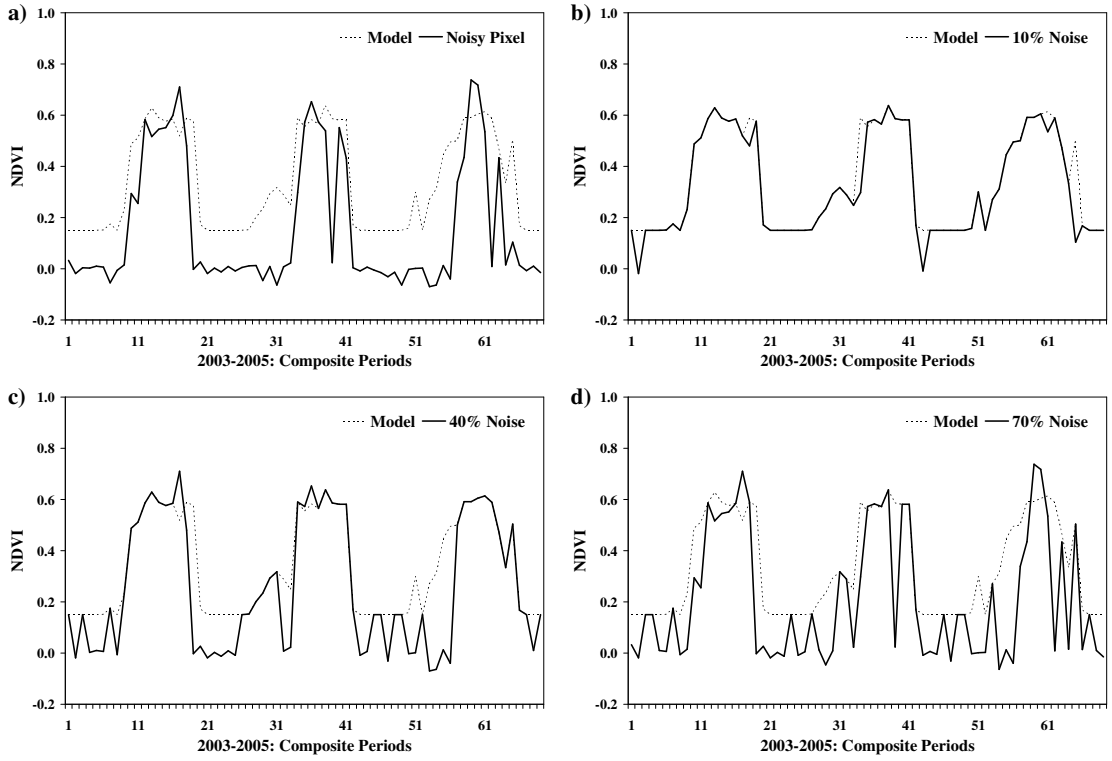

Figure A - 6: Montane model NDVI time series (dotted line), showing the a) selected noisy pixel time series, b) 10% introduced noise, b) 40% introduced noise and c) 70% introduced noise (solid lines).

APPENDIX B: IDL CODE USED IN IMPLEMENTING 4253H-TWICE, MVI AND ARMA3-ARMD5 FILTERS

B.1. IDL Code for the 4253H-Twice Filter

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;
;FUNCTION:     jk_4253hx2_singTS_v2.pro ( TO CALL: Result = jk_4253x2_sintgleTS('inTS') )
;
;Purpose:      To smooth a series of data values, using a set of running medians of
;              various widths (i.e. 4, 2, 5 then 3), followed by a Hanning filter (a
;              running average filter with weights '0.25, 0.5 and 0.25'), and then re-
;              roughing (applying the same smoothing procedure to the residuals and then
;               adding them to the smoothed data series.
;
;Arguments:    Name of array containing (NDVI) time series
;
;Returns:      Smoothed time series into variable specified
;              See: Velleman, P. 1980. Definition and Comparison of Robust Nonlinear Data
;              Smoothing Algorithms. In: Journal of the American Statistical Association,
;              Vol. 75(371): 609-615.
;
;Based on:     Velleman's (1980) 4253H, twice filter, and other descriptions of
;              the same filter
;              E.g. Van Dijk et al. (1987) applied it to vegetation index time series
;
;"Author":     This version: Jennifer Hird (MSc Candidate, Dept. Geogrpahy, U of Calgary)
;
;Last Updated: April 4, 2007
;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

FUNCTION jk_4253hx2_singTS_v2, inTS

IF ((TOTAL(inTS)) NE 0) THEN BEGIN

        ;---------- Query size of single time series
        nb = N_ELEMENTS(inTS)

        ;---------- If there are 7 or more elements in the array, then continue
        IF (nb GE 7) THEN BEGIN

                ;---------- Create float array and fill with data from the input array
                dataSeries = FLTARR(nb, /NOZERO)
                FOR i=0,nb-1 DO dataSeries[i] = inTS[i]*1.0

                ;---------- Create temporary arrays to store: results from 1st running
                ;median filter because the results will be offset..(see Velleman 1980);
                ;original time series, for computing residuals later, and for smoothed
                ;time series
                tempSeries = FLTARR (nb)
                smSeries = FLTARR(nb, /NOZERO)
                resSeries = FLTARR(nb, /NOZERO)

                ;       ---------- Fill variables with data from dataSeries for pixel [i,j]
                smSeries = dataSeries

                ;       ---------- Run '4' running median filter, store in temporary
                ;variable

                FOR k=2, nb-2 DO $
                        tempSeries[k] = MEDIAN(dataSeries[k-2:k+1],/EVEN)


                ;       ---------- Run '2' running median filter, using results from
```

```
;tempSeries
FOR k=2, nb-3 DO $
       smSeries[k] = MEDIAN(tempSeries[k:k+1],/EVEN)

;      ---------- Run '5' running median filter
tempSeries2 = smSeries
FOR k=2, nb-3 DO $
       smSeries[k] = MEDIAN(tempSeries2[k-2:k+2],/EVEN)


;      ---------- Run '3' running median filter
tempSeries3 = smSeries
FOR k=1, nb-2 DO $
       smSeries[k] = MEDIAN(tempSeries3[k-1:k+1],/EVEN)


;      ---------- Run Hanning filter (running average, with weights of
;0.25, 0.5 and 0.25), by creating a temporary vector to run 'TOTAL' on and
;calculate averages
tempSeries4 = smSeries
FOR k=1, nb-2 DO BEGIN
       tempVec   =   [tempSeries4[k-1],   tempSeries4[k],   tempSeries4[k],
tempSeries4[k+1]]
       smSeries[k] = (TOTAL(tempVec))*0.25
ENDFOR

;      ---------- Re-roughing: run all previous steps on residuals, than
;add to smoothed smSeries
resSeries = dataSeries - smSeries
FOR k=0, nb-1 DO BEGIN
       tempSeries[k] = 0.0
       tempSeries2[k] = 0.0
       tempSeries3[k] = 0.0
       tempSeries4[k] = 0.0
ENDFOR
tempVec = 0.0

;      ---------- Run the '4', '2', '5' and '3' median filters, then
;Hanning weighted running mean
FOR k=2, nb-2 DO $
       tempSeries[k] = MEDIAN(resSeries[k-2:k+1],/EVEN)
FOR k=2, nb-3 DO $
       resSeries[k] = MEDIAN(tempSeries[k:k+1],/EVEN)
tempSeries2 = resSeries
FOR k=2, nb-3 DO $
       resSeries[k] = MEDIAN(tempSeries2[k-2:k+2],/EVEN)
tempSeries3 = resSeries
FOR k=1, nb-2 DO $
       resSeries[k] = MEDIAN(tempSeries3[k-1:k+1],/EVEN)
tempSeries4 = resSeries
FOR k=1, nb-2 DO BEGIN
       tempVec   =   [tempSeries4[k-1],   tempSeries4[k],   tempSeries4[k],
tempSeries4[k+1]]
       resSeries[k] = (TOTAL(tempVec))*0.25
ENDFOR

;      ---------- Add smoothed residuals to original smoothed time series
smSeries = smSeries + resSeries

;---------- Prompt user to select or create a text file for output

outFile = DIALOG_PICKFILE (Title = 'Please select or create a text file
for output: ')

;---------- Open/create file specified by user for writing/output
OPENW, lun1, outFile, /GET_LUN

;---------- Print message to user in output file
;      Print message to user in output file
PRINTF, lun1, '*******************************************'
```

```
                PRINTF, lun1, 'Output of jk_4253hx2_singleTS.pro.'
                PRINTF, lun1, ' '
                PRINTF, lun1, 'Operates on a single time series; prints the original and'
                PRINTF, lun1, 'smoothed time series results, from the 4253Hx2 filter
(Velleman 1980).'
                PRINTF, lun1, 'The left column is the original time series NDVI values,
and '
                PRINTF, lun1, 'the right column is the new, smoothed NDVI values.'
                PRINTF, lun1, ' '
                PRINTF, lun1, '*********************************************'
                PRINTF, lun1, ' '
                PRINTF, lun1, ' '
                PRINTF, lun1, 'ELEMENT   ORIGINAL   SMOOTHED'
                PRINTF, lun1, '-------------------------------------'

                ;---------- Print output into text file
                FOR i=0, nb-1 DO PRINTF, lun1, i+1, ';', dataSeries[i], ';', smSeries[i]
                PRINTF, lun1, ' '
                PRINTF, lun1, '---------------------'
                PRINTF, lun1, 'END OF TIME SERIES.'
                PRINTF, lun1, ' '
                PRINTF, lun1, ' '
                PRINTF, lun1, 'END OF FILE.'
                PRINTF, lun1, '**************'

                ;---------- Print message to user, filename and successful completion
                PRINT, ' '
                PRINT, '*********************************************'
                PRINT, 'Output to File: ', outFile
                PRINT, ' '
                PRINT, '*********************************************'
                PRINT, '4253H, twice filtering completed successfully!'
                PRINT, ' '

                CLOSE, lun1

                ;---------- Return smSeries into variable specified by user when calling
                ;function
                RETURN, smSeries
        ENDIF ELSE BEGIN ;IF (nb GE 7) THEN BEGIN
                PRINT, ' '
                PRINT, '*************************************************'
                PRINT, 'Insufficient number of elements in time series array.'
                PRINT, '4253H, twice filtering NOT completed successfully.'
                PRINT, ' '
                RETURN, 0
        ENDELSE
ENDIF ELSE BEGIN ;IF ((TOTAL(inTS)) NE 0) THEN BEGIN
        PRINT, ' '
        PRINT, '*************************************************'
        PRINT, 'Invalid array variable.'
        PRINT, '4253H, twice filtering NOT completed successfully.'
        PRINT, ' '
        RETURN, 0
ENDELSE

END
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

B.2. IDL Code for the MVI Filter

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;
;FUNCTION jk_mviFilter_singleTS
;
;Author:       Ma and Veroustraete's (2006); received source code for the filter from
;              Mingguo Ma June 21, 2006 (written in C, translated by Jen Hird, MSc
;              Candidate, Department of Geography, U of Calgary)
;              URL:
;      http://individual.westgis.ac.cn/mmg/Content/Program/NSFC_2005/Reconstruct.cpp
;
;PURPOSE:      To smooth a (single) NDVI time series using Ma and Veroustraete's (2006)
Maximum
;              value iteration filter (calculates a multi-year average NDVI for each
;              composite/time period, then uses an iterative process to smooth the
;              time series
;
;PARAMETERS:   ndviTS - the single-dimension array containing the NDVI time series
;
;KEYWORDS:     SCALE_FACTOR - the scale factor by which the data would need to be divided
;              in order for it to be on the proper +1 to -1 NDVI scale
;
;Last Updated: April 11, 2007
;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

FUNCTION jk_mviFilter_singleTS, ndviTS, SCALE_FACTOR = sclfact
ndviTS = ndviTS
numElem = N_ELEMENTS(ndviTS)

IF ((TOTAL(ndviTS)) NE 0) THEN BEGIN

;---------- If keyworkd SCALE_FACTOR is set (i.e. NDVI values have been scaled and need
;---------- to be rescaled back to normal -1 to +1 scale) then rescale them by scale
;---------- factor given. ;E.g. for MODIS, NDVI are scaled by 10,000, and to return them
;---------- to normal NDVI values,divide by 10,000.
;-------------------------------------------------------------------------------------

IF (sclfact NE 0) THEN ndviTS = ndviTS/sclfact
origTS = ndviTS

;---------- Create variable 'myAvg' to hold the multi-year average
myAvg = 0.0


;---------- Create variable 'threshold' to hold the %10 of multi-year average for each
;---------- composite period in a year
threshold = 0.0

;---------- Calculate multi-year average.  In a series of images this is done for each
;---------- pixel - it calculates the multi-year average NDVI for that pixel.  But here,
;---------- the program just deals with one time series (i.e. one 'pixel').
myAvg = (TOTAL(ndviTS)) / numElem
threshold = myAvg * 0.1

;---------- Iterative mean-value filtering process, using the threshold
;---------- calculated above; uses a 1-D array to contain a temporary average
;---------- time series, calculated for each pixel
;-----------------------------------------------------------------

;---------- Create 1-D floating point array to contain a time series
avgTS = FLTARR(numElem, /NOZERO)

;---------- Iterative mean-value filtering process, repeats until (id EQ -1)
REPEAT BEGIN

       ;---------- Fill array 'avgTS' with averages of original NDVI values (from one
       ;---------- before and one after)
```

```
    FOR i=1,numElem-2 DO BEGIN
       avgTS[i] = (ndviTS[i-1] + ndviTS[i+1])*0.5
    ENDFOR

    ;---------- Create variable 'dif' to hold absolute difference between
    ;---------- average values in 'avgTS' and original NDVI values in 'ndviTS'
    dif = 0.0

    ;---------- Create variable 'id' to hold the element (composite period) with the
    ;---------- biggest difference in 'dif'
    id = -1

    ;---------- For the second element to the second last do the following:
    FOR j=1, numElem-2 DO BEGIN

       ;---------- If the absolute difference between avgTS and ndviTS for the 'k'
       ;---------- composite period/element is greater than the threshold  greater
       ;---------- than 'dif' then replace the ndviTS with the avgTS
       ;---------- value; set 'dif' and 'k' and then repeat with next element
        IF ((ABS(avgTS[j] - ndviTS[j])) GT threshold) && $
        ((ABS(avgTS[j] - ndviTS[j])) GT dif) THEN BEGIN
              dif = ABS(avgTS[j] - ndviTS[j])
              id = j
        ENDIF
    ENDFOR
    IF (id GT -1) THEN ndviTS[id] = avgTS[id]
ENDREP UNTIL (id EQ -1)

;---------- Prompt user to select/create a textfile for output
outFile = DIALOG_PICKFILE (Title = 'Please select or create a text file for output: ')

;---------- Open/create file specified by user for writing/output
OPENW, lun1, outFile, /GET_LUN

       ;---------- Print message to user in output file
       ;       Print message to user in output file
       PRINTF, lun1, '****************************************'
       PRINTF, lun1, 'Output of jk_mvifilter_singleTS.pro.'
       PRINTF, lun1, ' '
       PRINTF, lun1, 'Operates on a single time series, prints the original and'
       PRINTF, lun1, 'smoothed time series results, from the mean-value iteration (MVI)'
       PRINTF, lun1, 'filter (Ma and Veroustraete 2006).
       PRINTF, lun1, ' '
       PRINTF, lun1, 'The left column is the original time series NDVI values, and the'
       PRINTF, lun1, 'the right column is the new, smoothed NDVI values.'
       PRINTF, lun1, ' '
       PRINTF, lun1, '****************************************'
       PRINTF, lun1, ' '
       PRINTF, lun1, ' '
       PRINTF, lun1, 'ELEMENT   ORIGINAL   SMOOTHED'
       PRINTF, lun1, '-------------------------------------'

       ;---------- Print output into text file
       FOR i=0, numElem-1 DO PRINTF, lun1, i+1, ';', origTS[i], ';', ndviTS[i]
              PRINTF, lun1, ' '
              PRINTF, lun1, '---------------------'
              PRINTF, lun1, 'END OF TIME SERIES.'
              PRINTF, lun1, ' '
              PRINTF, lun1, ' '
              PRINTF, lun1, 'END OF FILE.'
              PRINTF, lun1, '**************'

              ;---------- Print message to user, filename and successful completion
              PRINT, ' '
              PRINT, '********************************************'
              PRINT, 'Output to File: ', outFile
              PRINT, ' '
              PRINT, '********************************************'
              PRINT, 'MVI filtering completed successfully!'
              PRINT, ' '
```

```
                CLOSE, lun1

                ;---------- Return smSeries into variable specified by user when calling
                ;---------- function
                RETURN, ndviTS
ENDIF ELSE BEGIN ;IF ((TOTAL(inTS)) NE 0) THEN BEGIN
        PRINT, ' '
        PRINT, '**********************************************'
        PRINT, 'Invalid array variable.'
        PRINT, 'MVI filtering NOT completed successfully.'
        PRINT, ' '
        RETURN, 0
ENDELSE

END

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

B.3. IDL Code for the ARMD3-ARMA5 Filter

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;
;PROCEDURE:    jk_armd3arma5.pro
;
;PURPOSE:      To smooth a noisy time series by applying an autoregresive moving
;              median (armd) with a 3-element window, followed by an autoregressive
;              moving maximum (armm) with a five-element window
;
;*NOTE:        This was designed using 10-day AVHRR composite NDVI data; and is
;              applied to 16-day MODIS composite NDVI data
;
;ARGUMENT:     'inTS' - array containing single NDVI time series
;
;OUTPUT:       Prints the original and smoothed time series into a text file
;
;AUTHOR:       Based on unpublished work by Hall-Beyer & Filipova-Racheva, written
;              in IDL by Jen Hird (M.Sc. Candidate, Dept. Geography, U of Calgary)
;
;Last Updated: April 3, 2007
;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

PRO jk_armd3arma5, inTS
;---------- If sum of elements in time series is not zero, do the following...
IF ((TOTAL(inTS)) NE 0) THEN BEGIN

        ;---------- Query size of single time series
        nb = N_ELEMENTS(inTS)

        ;---------- If there are 7 or more elements in the array, then continue
        IF (nb GE 7) THEN BEGIN ;***COME BACK TO THIS - RIGHT NO. ELEMENTS?***

                ;---------- Create a floating point array 'smTS' for the smoothed time
                ;---------- series, and fill it with the values from 'inTS'
                smTS = FLTARR(nb, /NOZERO)
                FOR i=0,nb-1 DO smTS[i] = inTS[i]*1.0

                ;---------- Run a 3-element running median filter on smTS
                FOR j=1, nb-2 DO smTS[j] = MEDIAN(smTS[j-1:j+1], /EVEN)

                ;---------- Run a 5-element running average filter on smTS; create
                ;---------- temporary array, for each processable element fill temporary
                ;---------- array with 5 surrounding values, calculate average and place
                ;---------- back into smTS
                tempTS = FLTARR(5, /NOZERO)
```

```
		FOR k=2, nb-3 DO BEGIN
			tempTS = [smTS[k-2], smTS[k-1], smTS[k], smTS[k+1], smTS[k+2]]
			smTS[k] = (TOTAL(tempTS))*0.20
		ENDFOR

		;---------- Prompt user to select or create a text file for output
		outFile = DIALOG_PICKFILE (Title = 'Please select or create a text file 
for output: ')
		;---------- Open/create file specified by user for writing/output
		OPENW, lun1, outFile, /GET_LUN

		;---------- Print message to user in output file
		;---------- Print message to user in output file
		PRINTF, lun1, '********************************************'
		PRINTF, lun1, 'Output of jk_armd3arma5.pro.'
		PRINTF, lun1, ' '
		PRINTF, lun1, 'Operates on a single time series, prints the original and'
		PRINTF, lun1, 'smoothed time series results, from the ARMD3ARMA5 filter '
		PRINTF, lun1, '(Hall-Beyer and Filipova-Racheva, unpublished).'
		PRINTF, lun1, ' '
		PRINTF, lun1, 'The left column is the element number, the middle column is 
the'
		PRINTF, lun1, 'original time series NDVI value, and the right column is 
the new,'
		PRINTF, lun1, 'smoothed NDVI values.'
		PRINTF, lun1, '********************************************'
		PRINTF, lun1, ' '
		PRINTF, lun1, ' '
		PRINTF, lun1, 'ELEMENT       ORIGINAL       SMOOTHED'
		PRINTF, lun1, '-------------------------------------'

		;---------- Print output into text file
		FOR i=0, nb-1 DO PRINTF, lun1, i+1, ';', inTS[i], ';',  smTS[i]
		PRINTF, lun1, ' '
		PRINTF, lun1, '-------------------------------------'
		PRINTF, lun1, 'END OF TIME SERIES.'
		PRINTF, lun1, ' '
		PRINTF, lun1, ' '
		PRINTF, lun1, 'END OF FILE.'
		PRINTF, lun1, '*************'

		;---------- Print message to user, filename and successful completion
		PRINT, ' '
		PRINT, '*********************************************'
		PRINT, 'Output to File: ', outFile
		PRINT, ' '
		PRINT, '*********************************************'
		PRINT, 'ARMD3ARMA5 filtering completed successfully!'
		PRINT, ' '
		CLOSE, lun1

	;---------- If number of elements is too few, do the following...
	ENDIF ELSE BEGIN ;IF (nb GE 7) THEN BEGIN
		PRINT, ' '
		PRINT, 'Not enough elements in array for smoothing, '
		PRINT, ' '
		PRINT, '*********************************************'
		PRINT, 'ARMD3ARMA5 filtering not successful.'
	ENDELSE
;---------- If sum of elements in time series IS zero, do the following...
ENDIF ELSE BEGIN ;IF ((TOTAL(inTS)) NE 0) THEN BEGIN
	PRINT, ' '
	PRINT, 'Sum of array equals zero (i.e. no data), '
	PRINT, ' '
	PRINT, '*********************************************'
	PRINT, 'ARMD3ARMA5 filtering not successful.'
ENDELSE
END
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

APPENDIX C: APPLICATION OF THE SIX NOISE REDUCTION TECHNIQUES TO THE NOISY NDVI TIME SERIES

C.1. Grassland Time Series

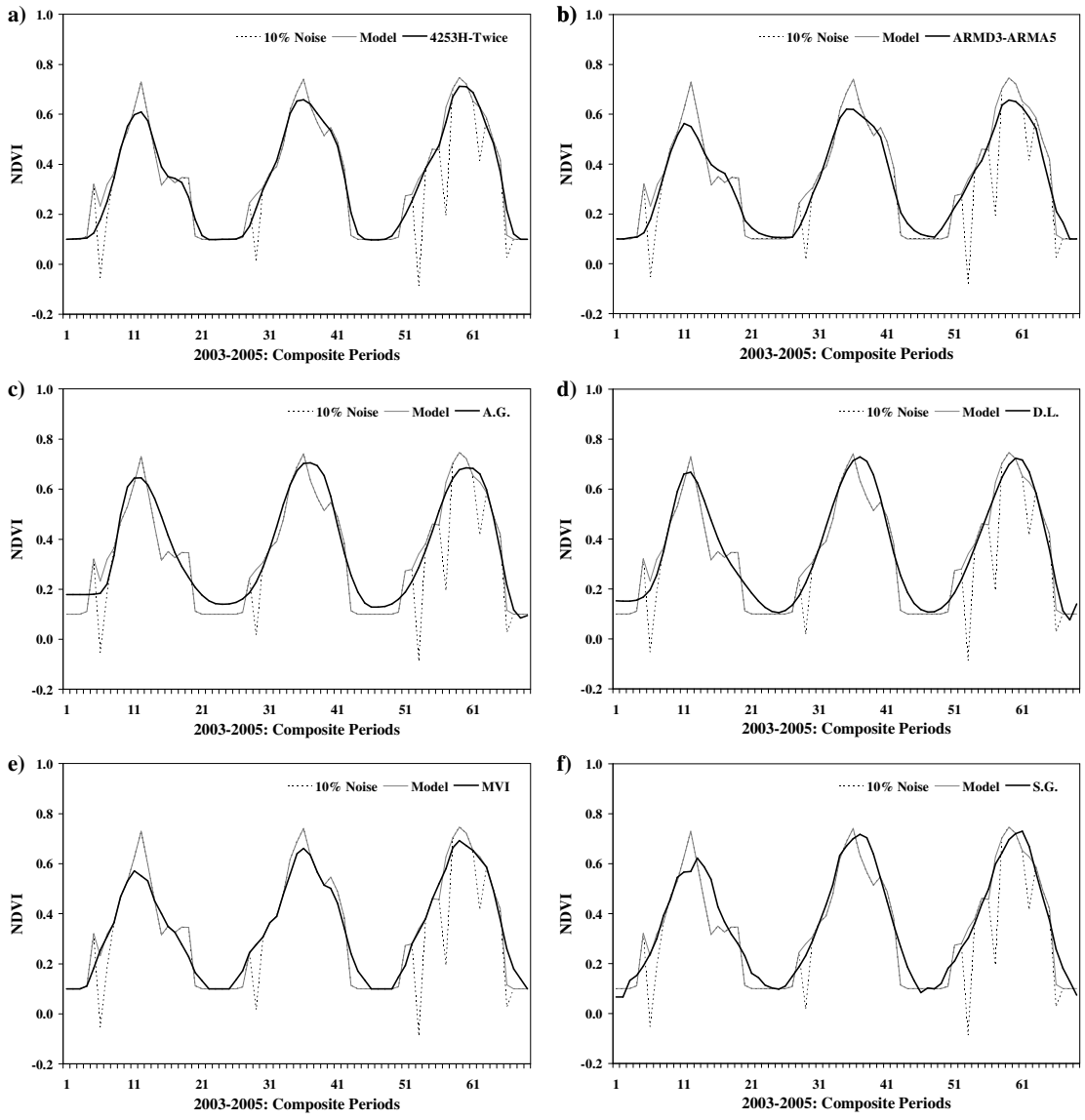

Figure C - 1: Grassland NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

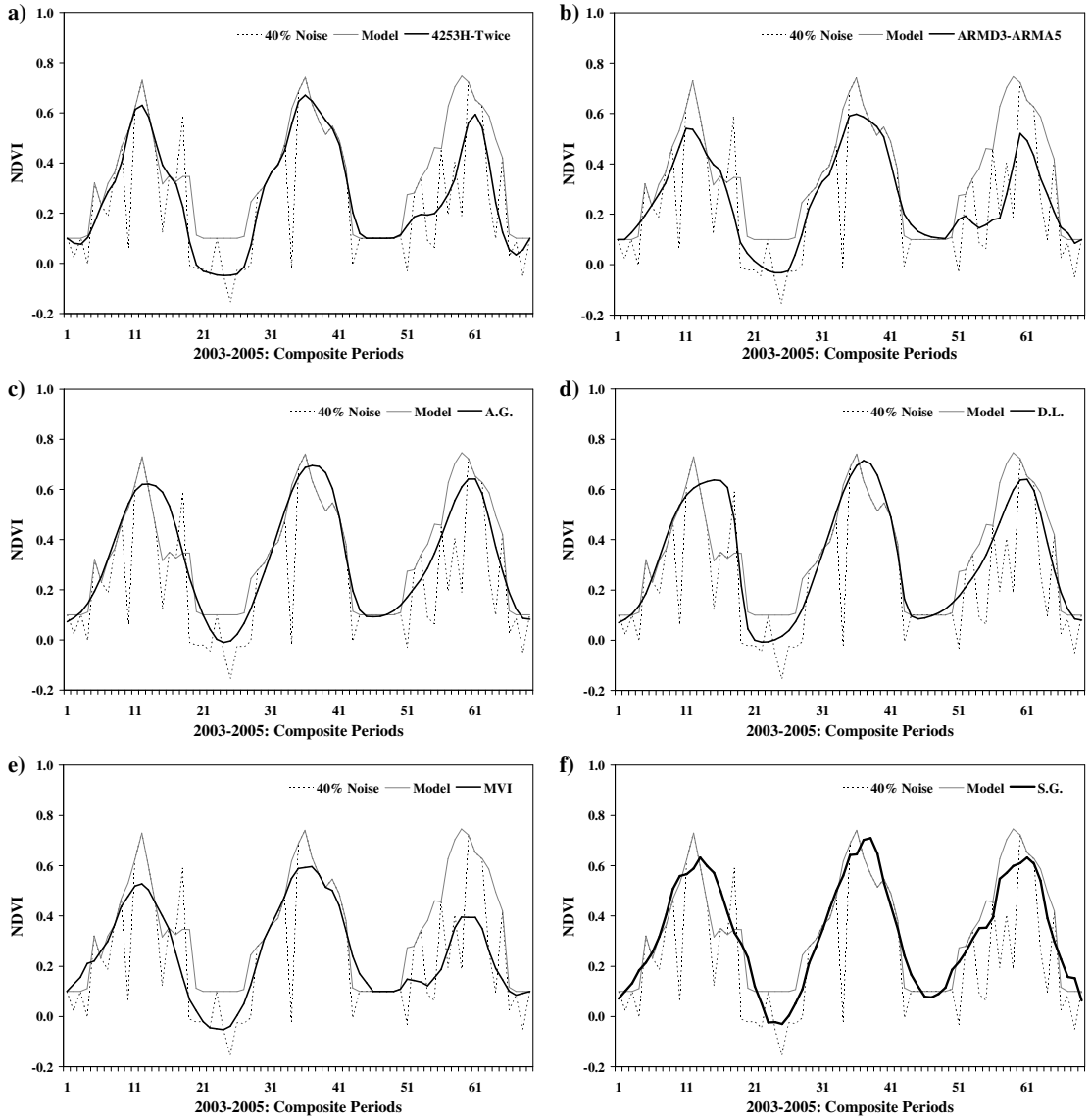

Figure C - 2: Grassland NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

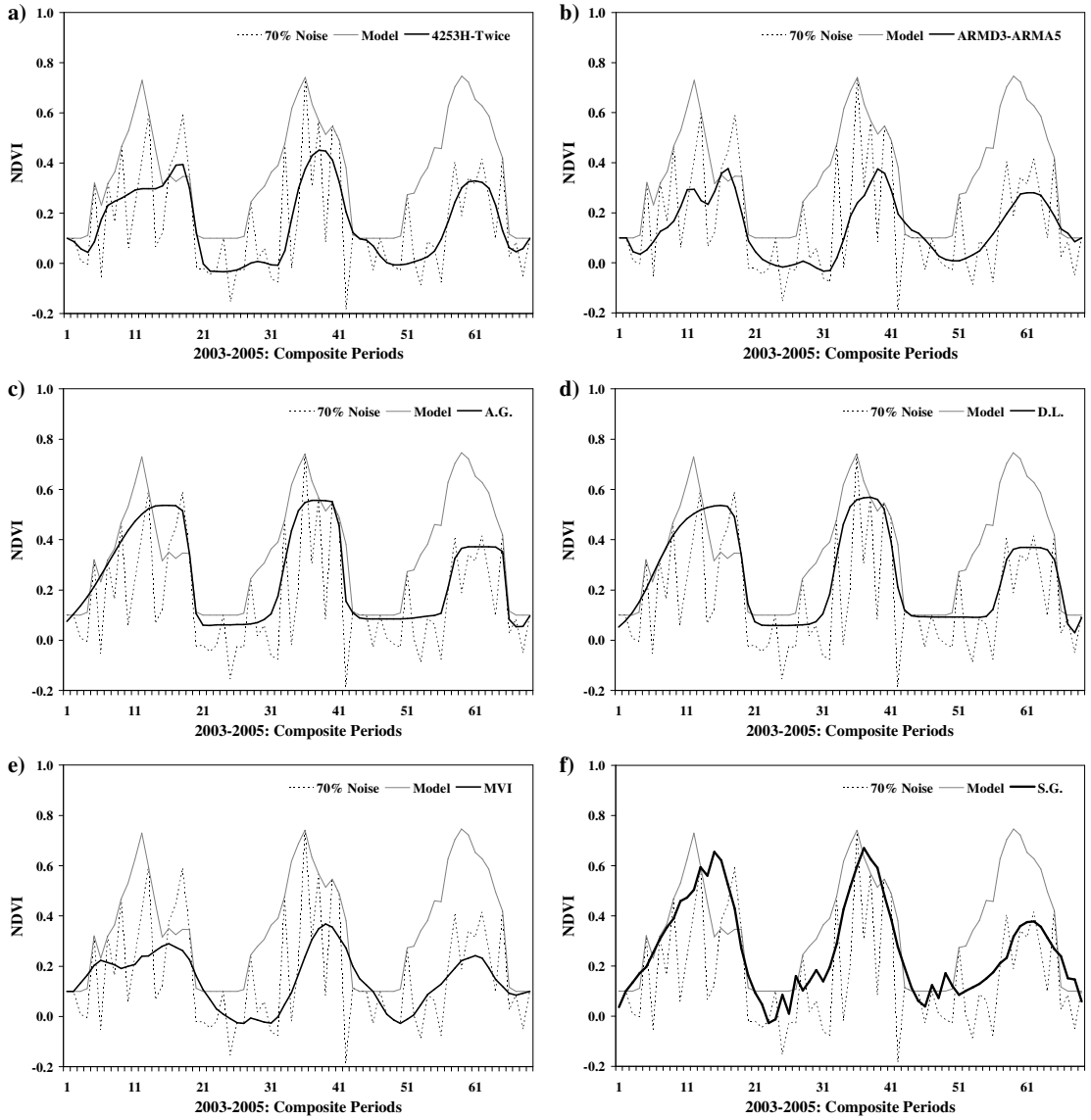

Figure C - 3: Grassland NDVI time series, showing the original model (grey line), the model with 70% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

C.2. Parkland Time Series

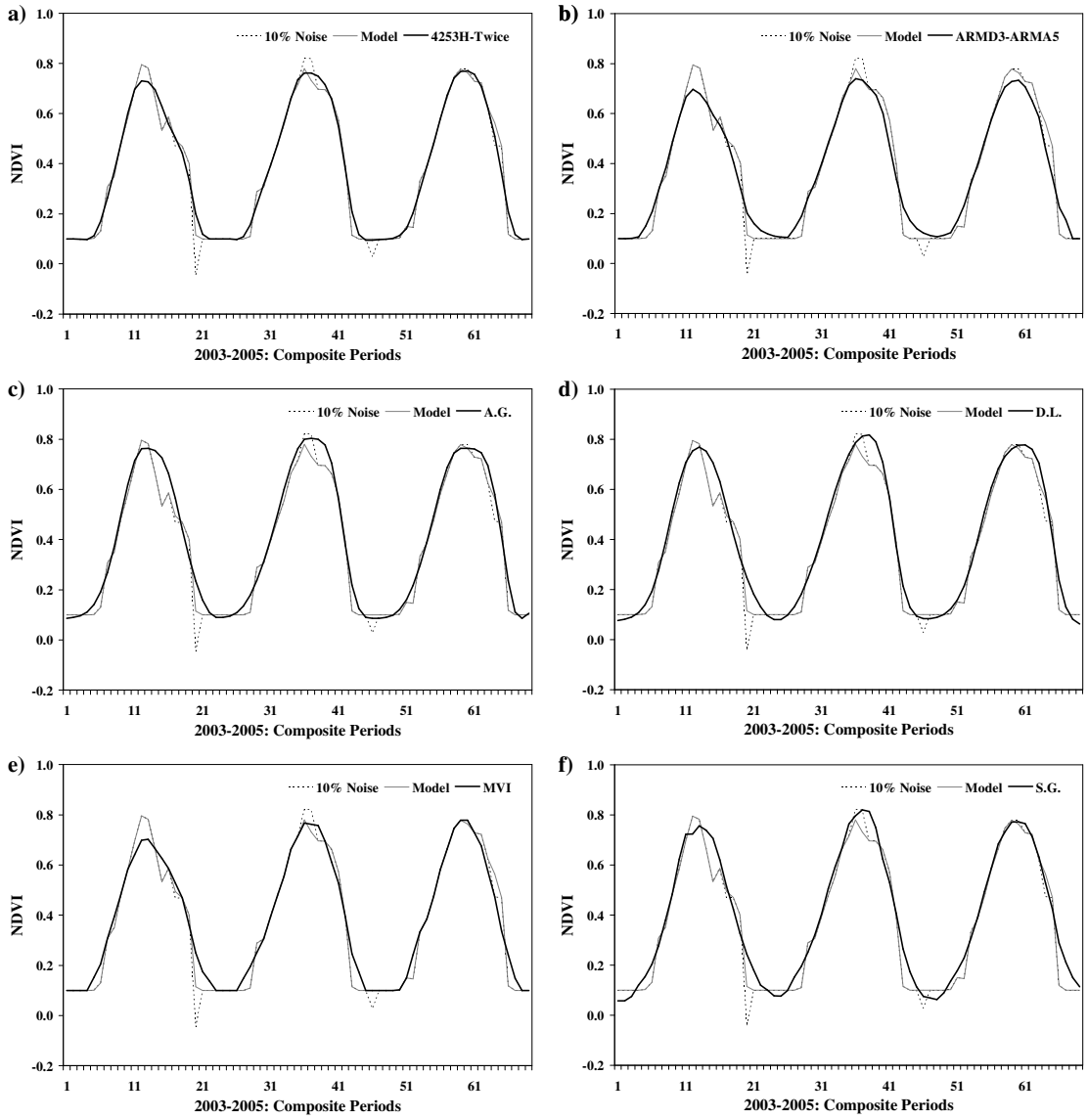

Figure C - 4: Parkland NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

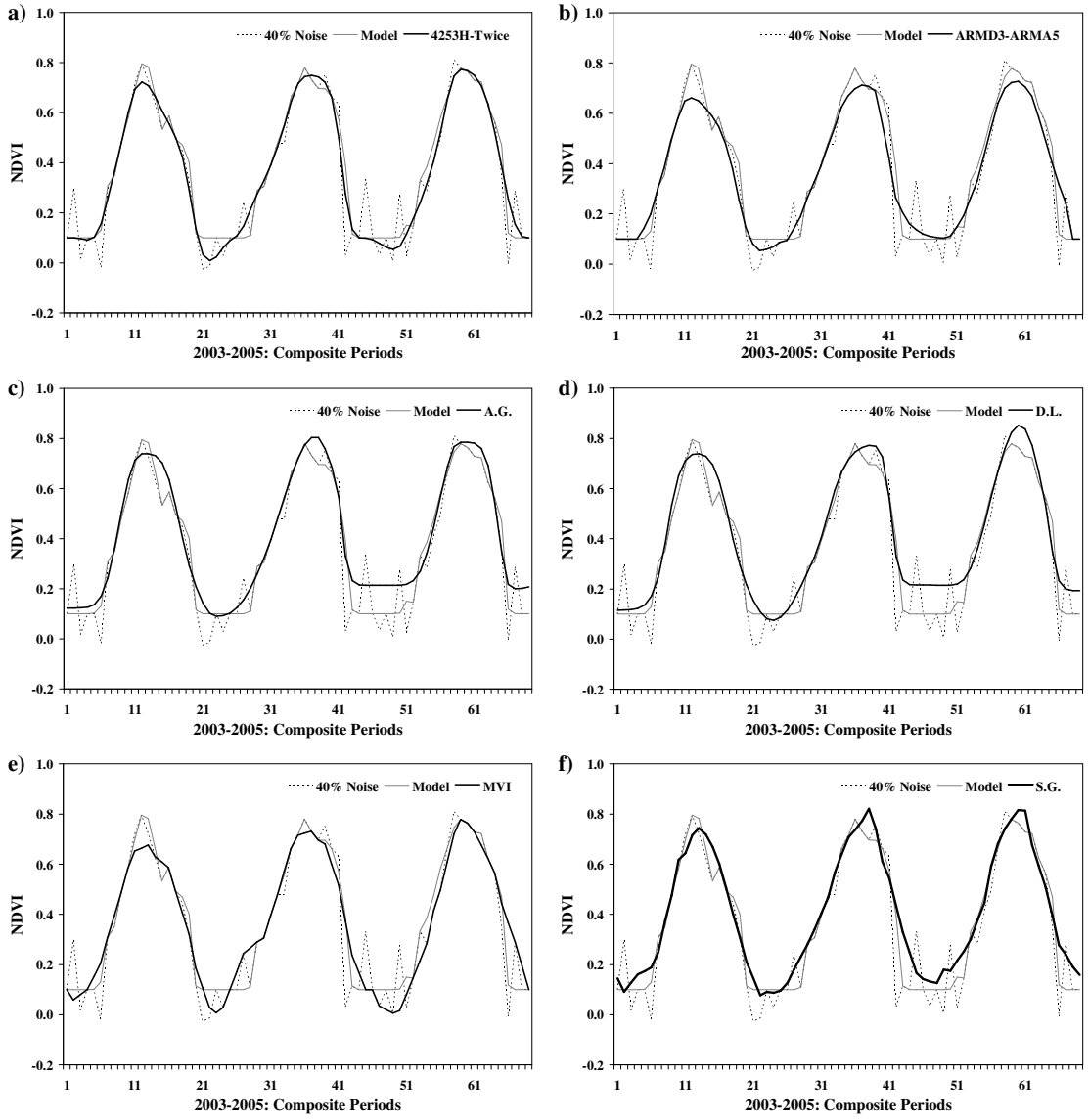

Figure C - 5: Parkland NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

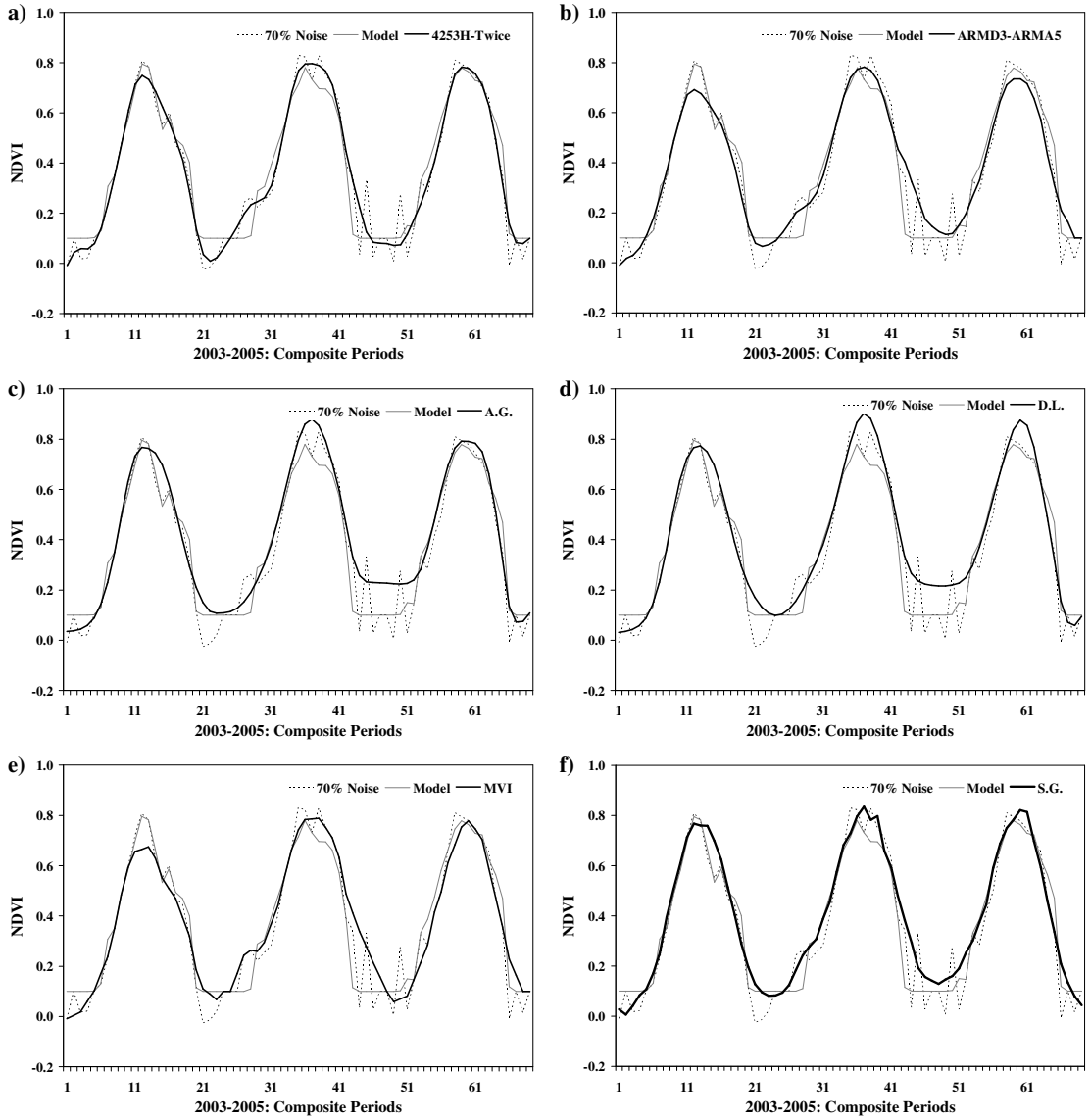

Figure C - 6: Parkland NDVI time series, showing the original model (grey line), the model with 70% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

C.3. Boreal Time Series

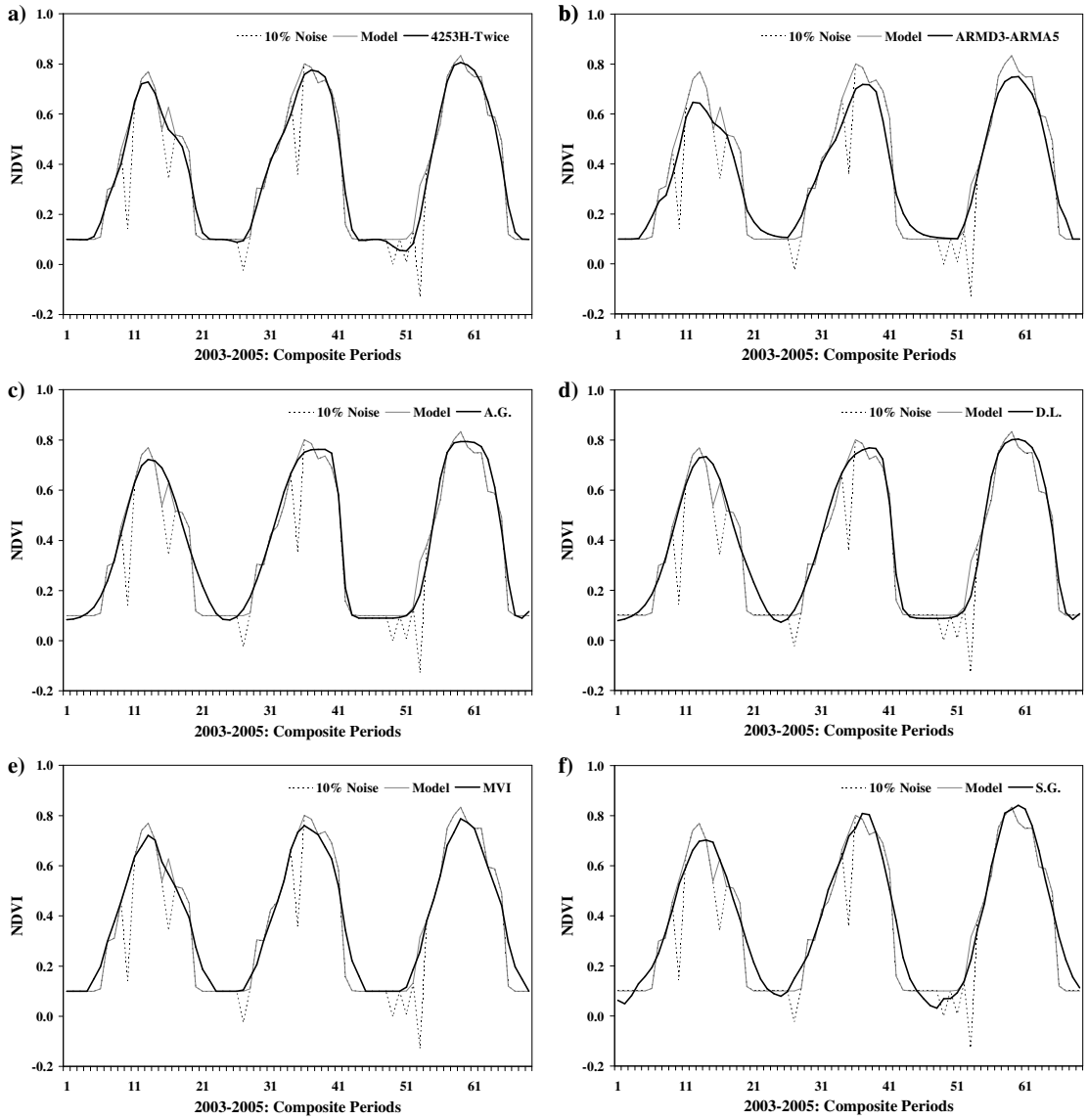

Figure C - 7: Boreal NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

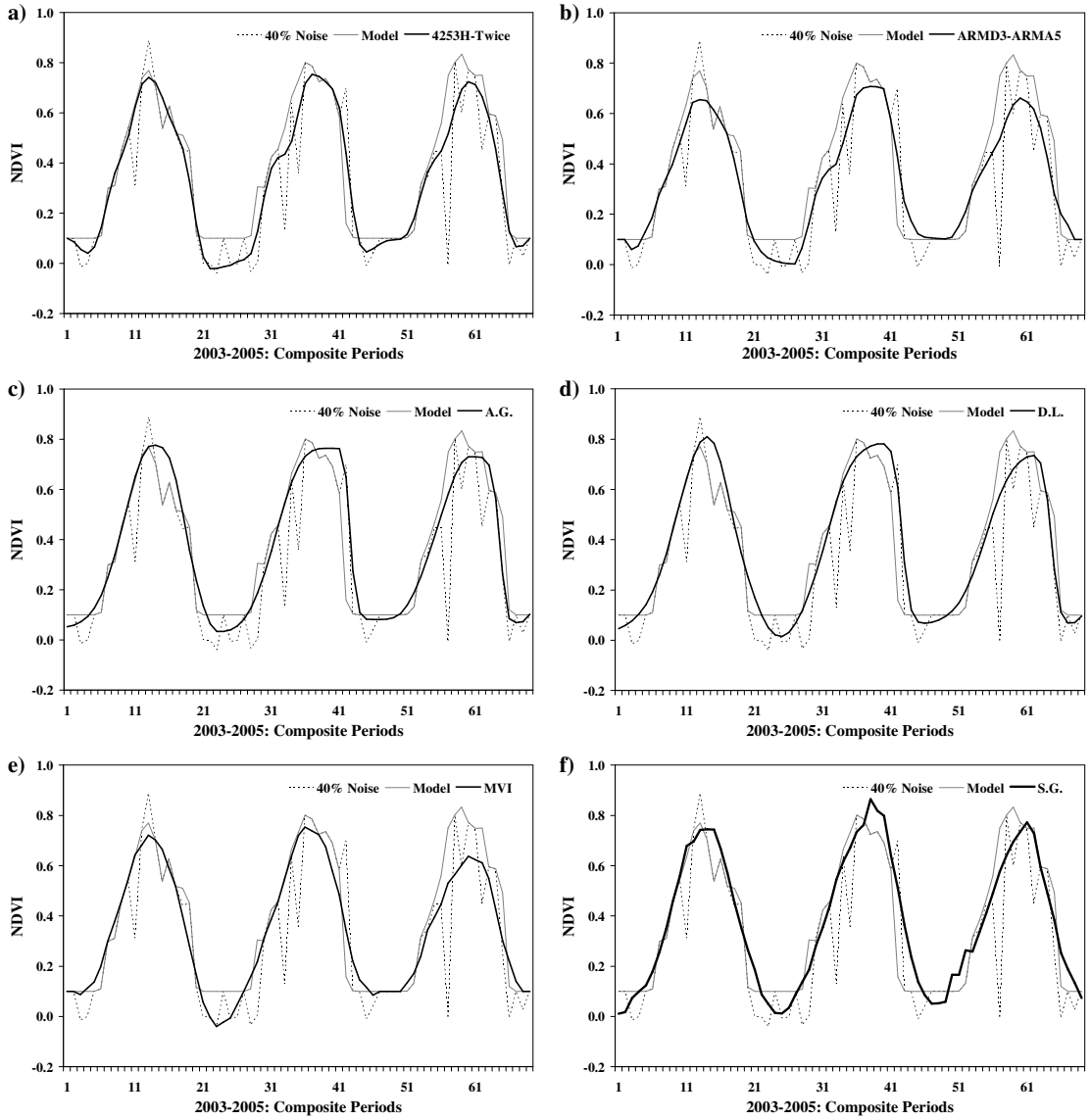

Figure C - 8: Boreal NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

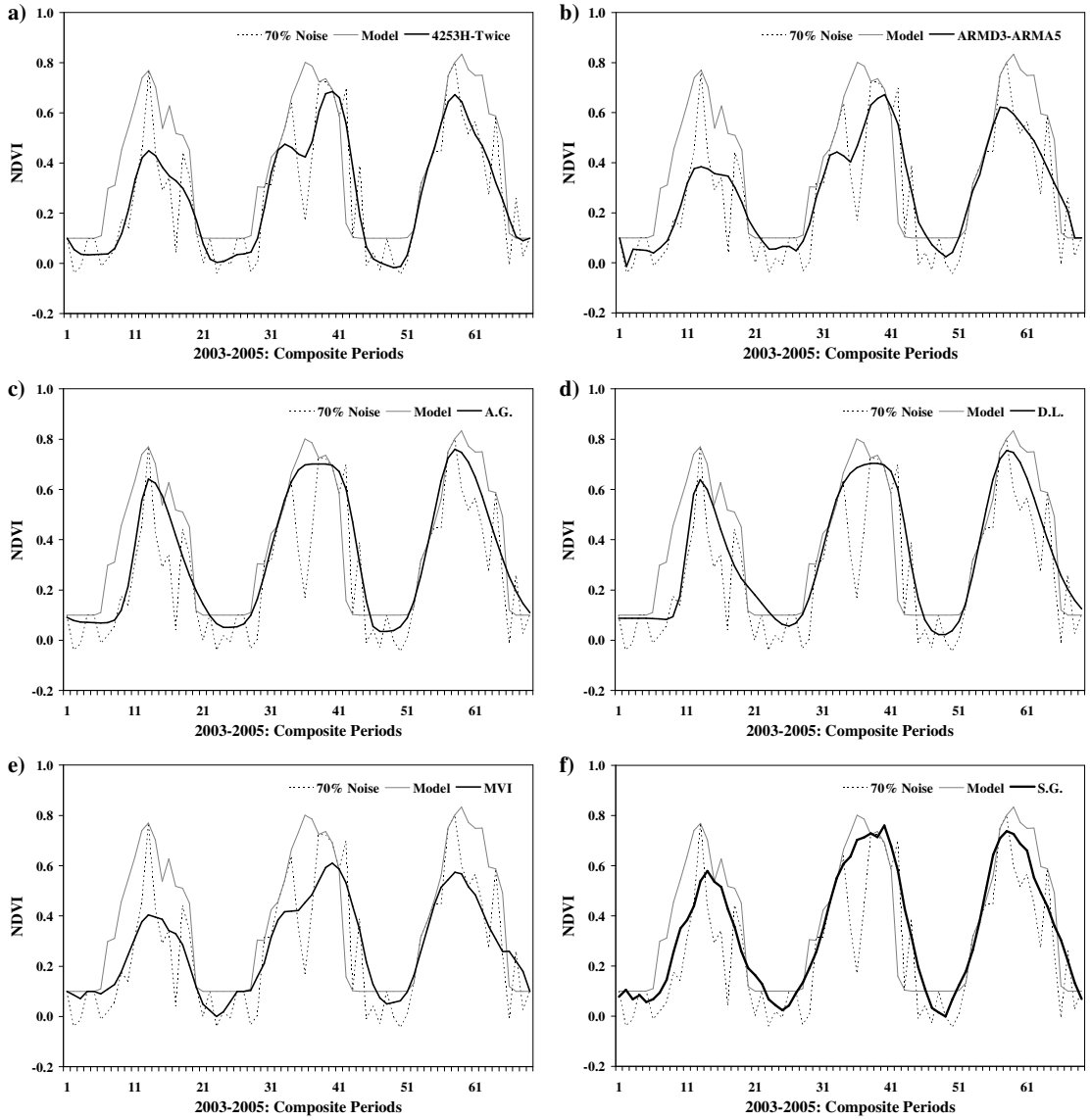

Figure C - 9: Boreal NDVI time series, showing the original model (grey line), the model with 70% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

C.4. Lower Foothills Time Series

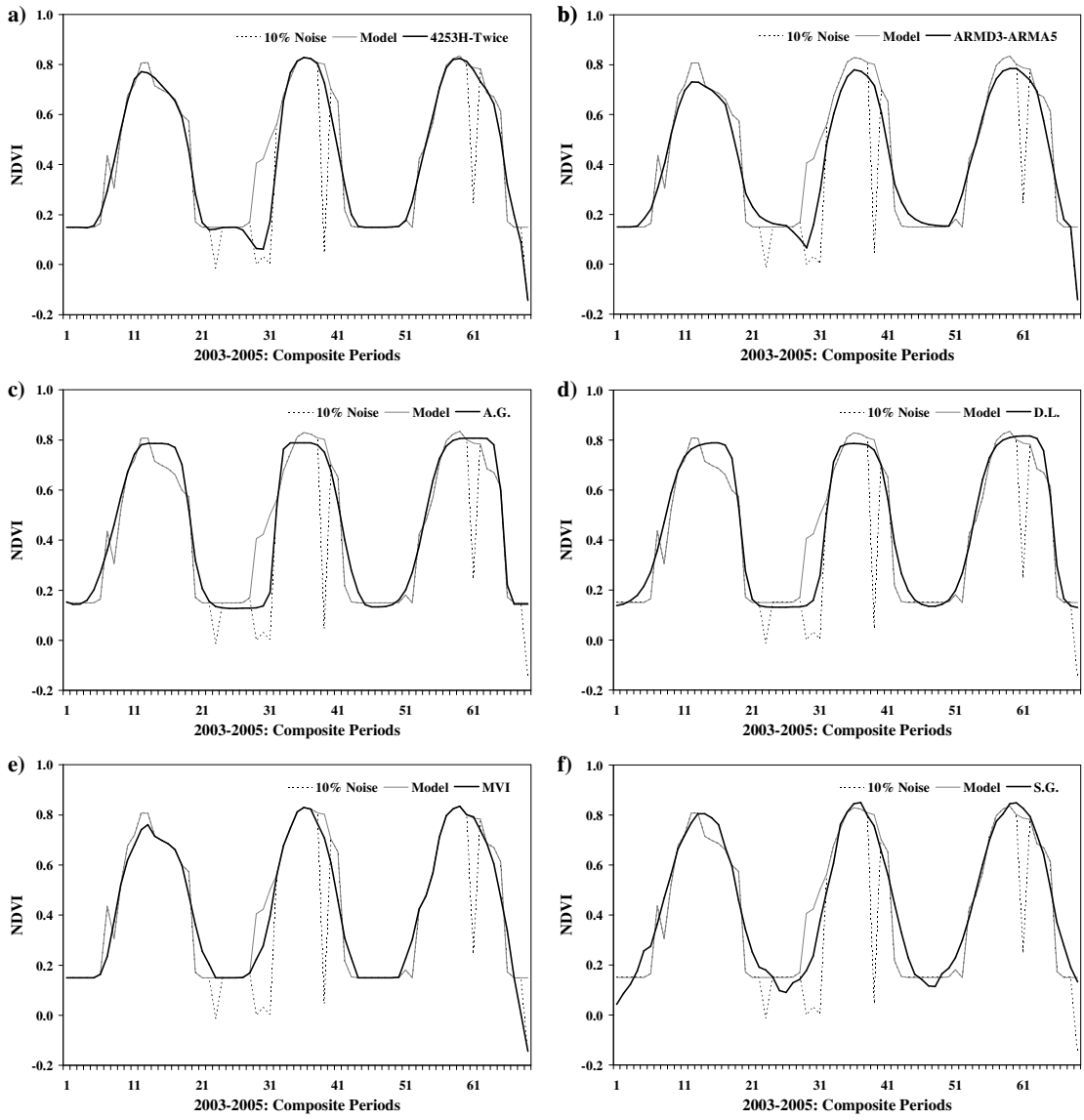

Figure C - 10: Lower Foothills NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

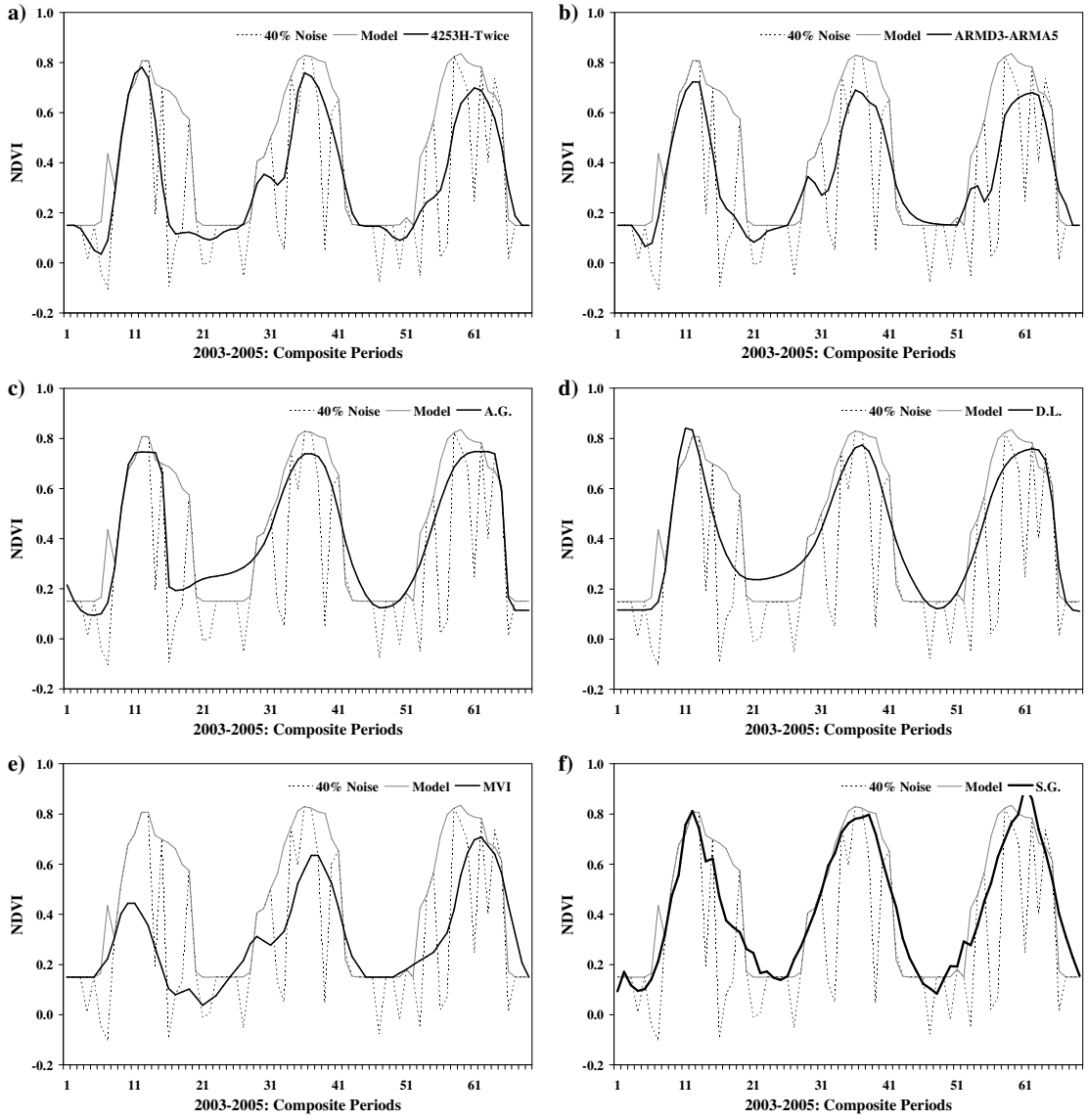

Figure C - 11: Lower Foothills NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

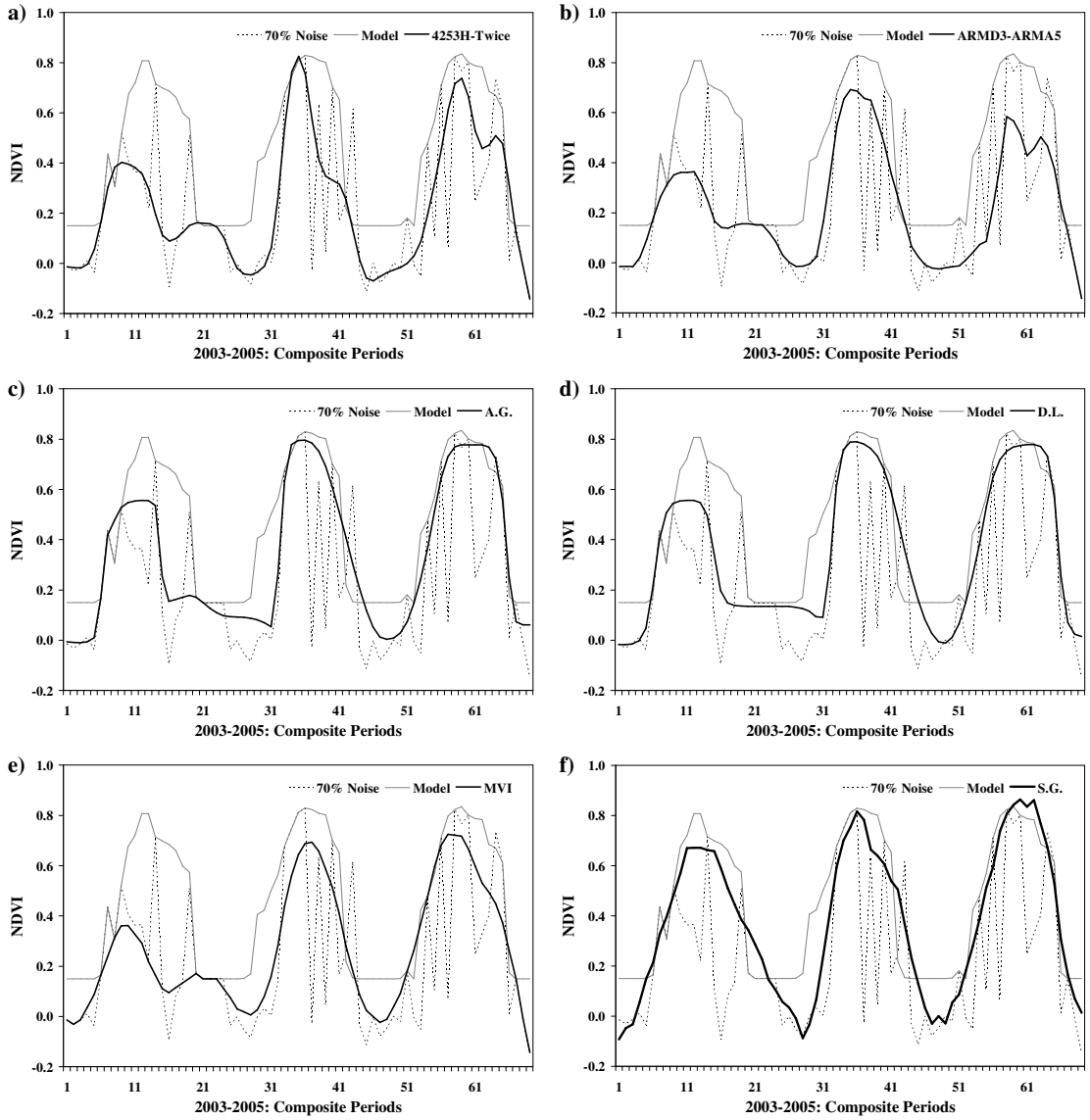

Figure C - 12: Lower Foothills NDVI time series, showing the original model (grey line), the model with 70% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

C.5. Montane Time Series

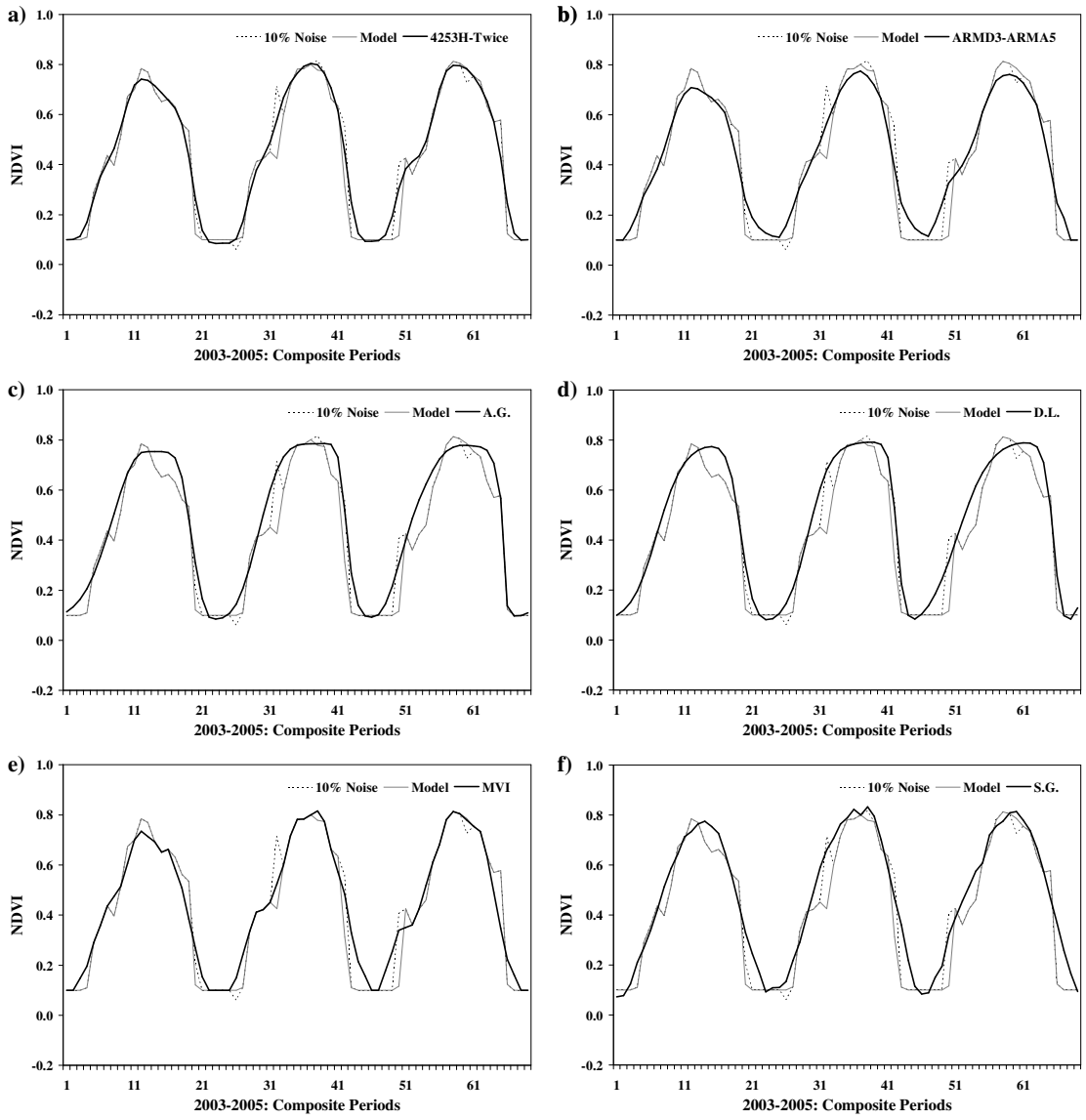

Figure C - 13: Montane NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

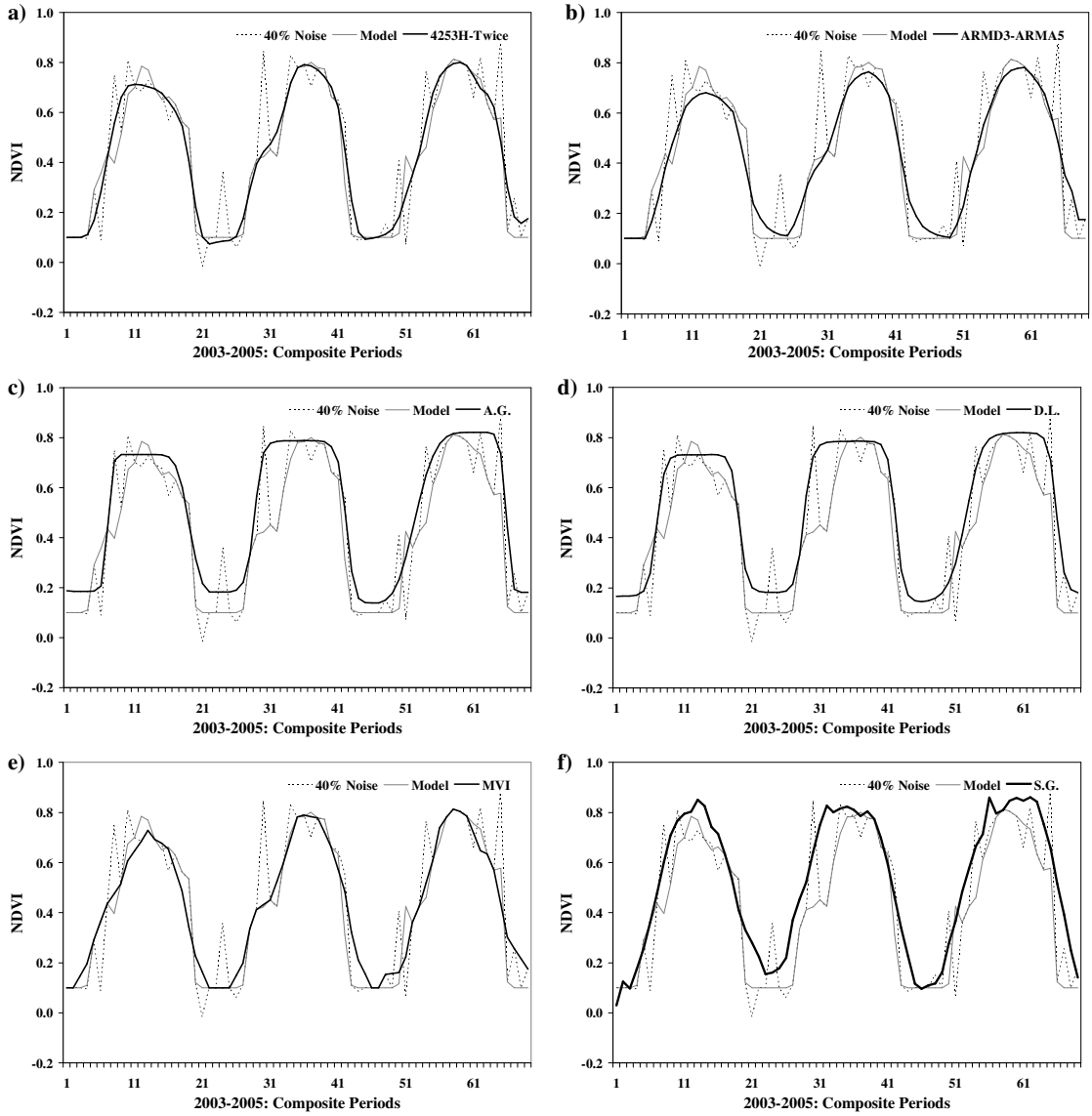

Figure C - 14: Montane NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

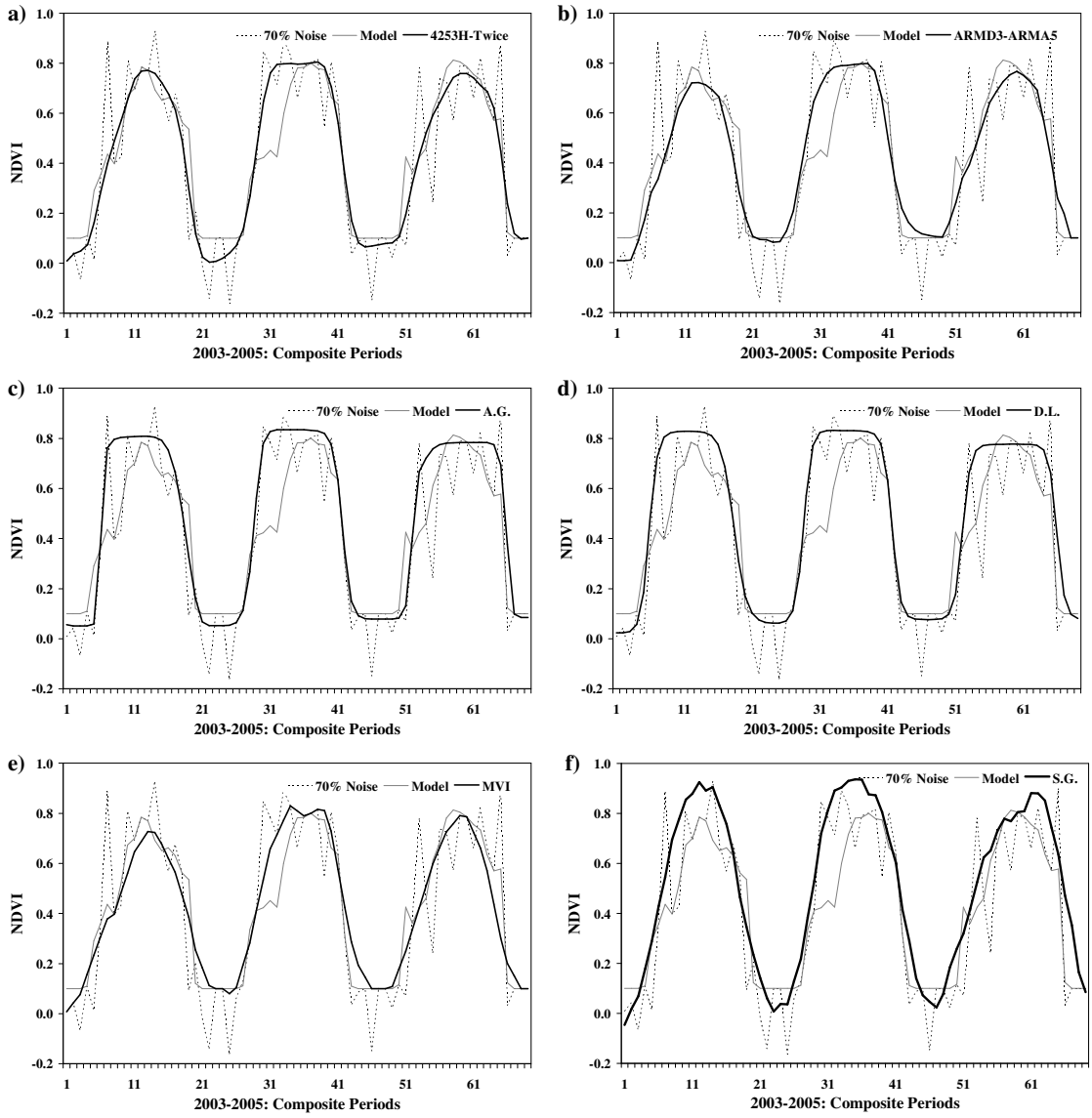

Figure C - 15: Montane NDVI time series, showing the original model (grey line), the model with 70% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

C.6. Alpine Time Series

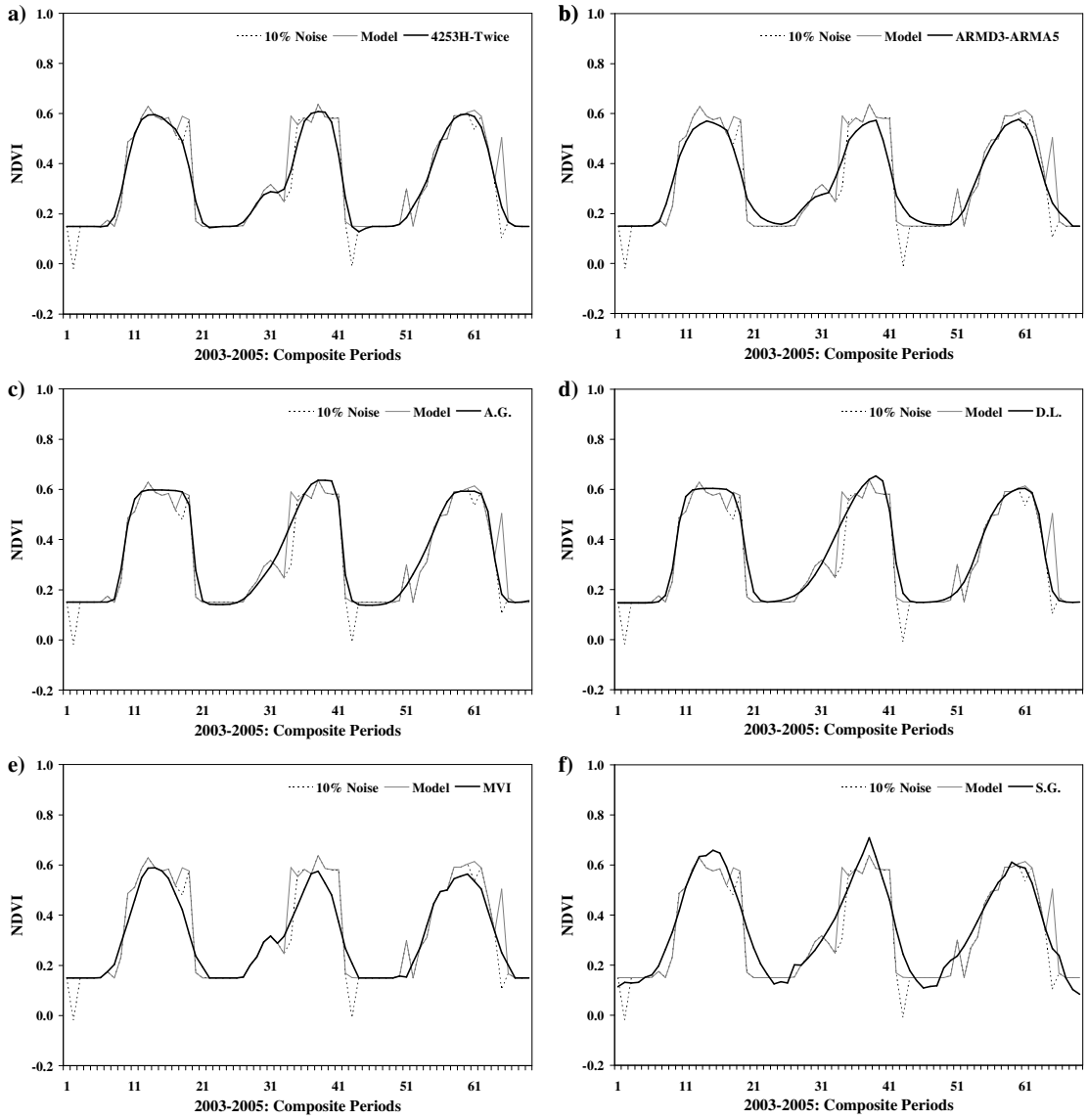

Figure C - 16: Alpine NDVI time series, showing the original model (grey line), the model with 10% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

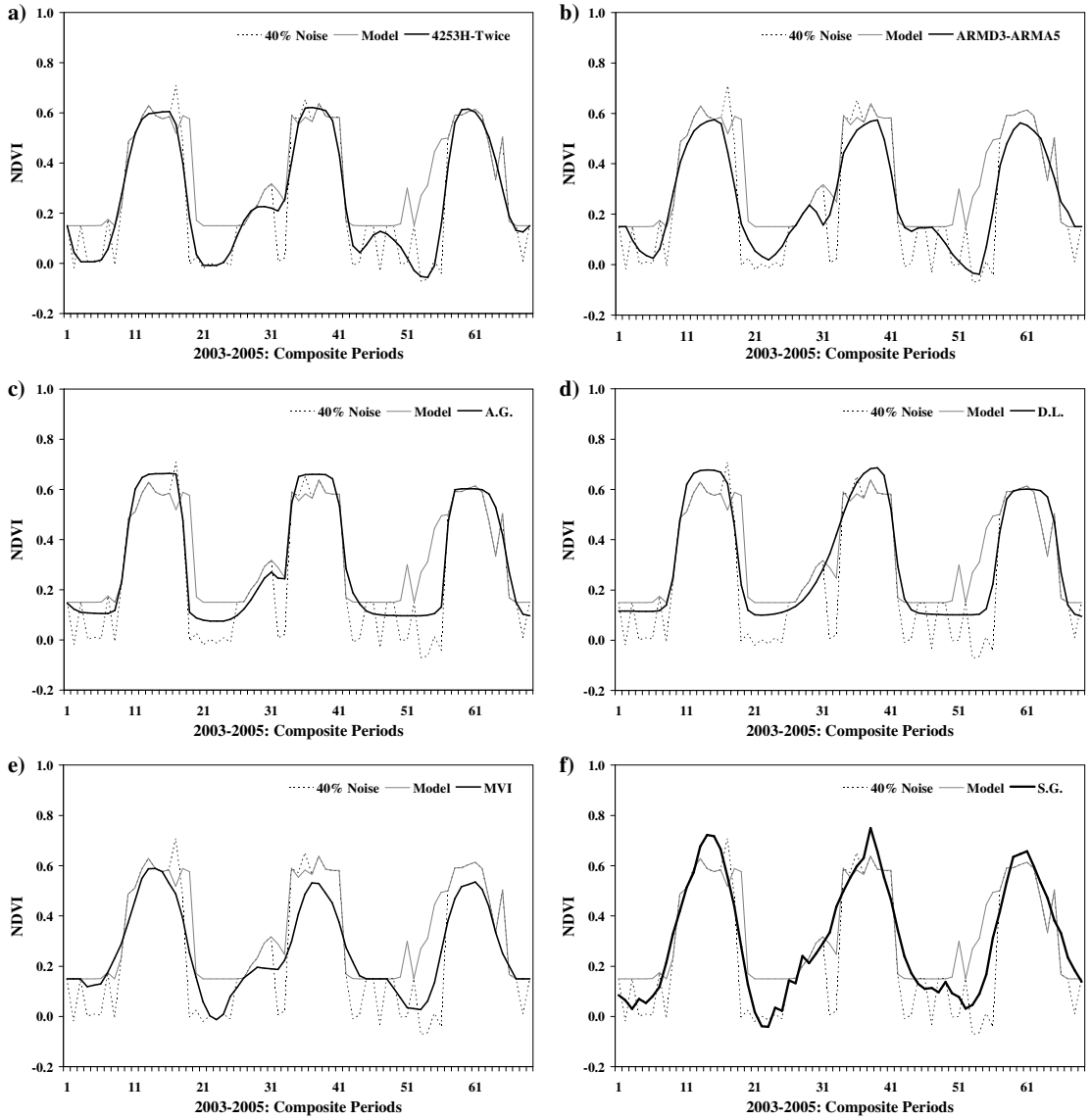

Figure C - 17: Alpine NDVI time series, showing the original model (grey line), the model with 40% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

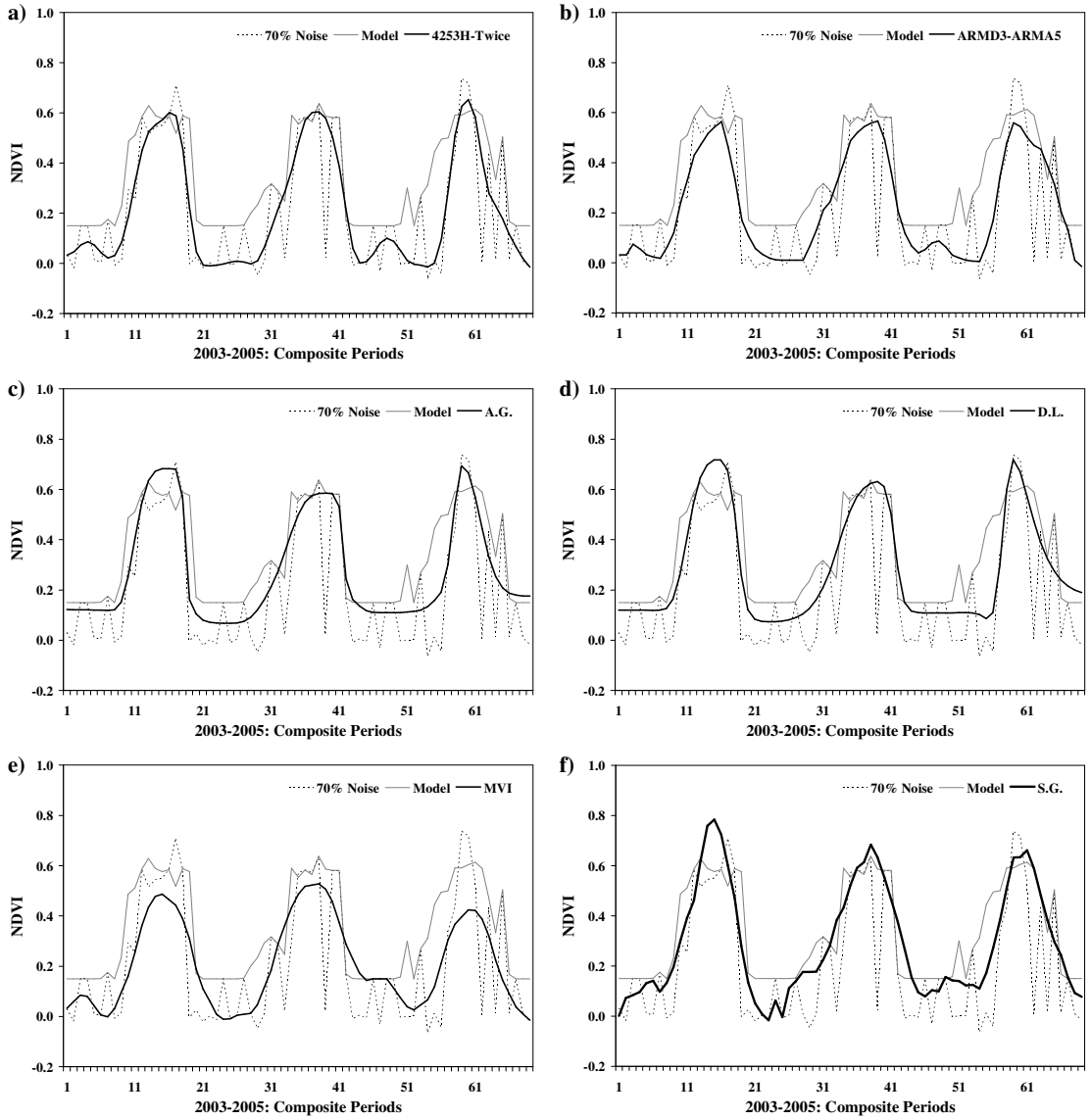

Figure C - 18: Alpine NDVI time series, showing the original model (grey line), the model with 710% introduced noise (dotted line), and application (black lines) of the a) 4253H-Twice filter, b) ARMD-ARMA5 filter, c) Asymmetric Gaussian function-fitting, d) Double Logistic function-fitting, e) MVI filter and f) Savitzky-Golay filter.

APPENDIX D: RMSE AND METRIC RESULTS

D.1. RMSE Calculation Results

Table D - 1: Results of the RMSE calculations for the six land cover types, three noise levels and application of the seven candidate noise reduction strategies.

	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Grassland							
10	0.046	0.060	0.066	0.062	0.052	0.066	0.092
40	0.128	0.143	0.085	0.100	0.153	0.084	0.185
70	0.240	0.266	0.170	0.168	0.269	0.173	0.265
Parkland							
10	0.035	0.050	0.050	0.052	0.049	0.057	0.027
40	0.048	0.062	0.067	0.069	0.071	0.067	0.083
70	0.062	0.074	0.076	0.081	0.082	0.071	0.079
Boreal							
10	0.046	0.066	0.053	0.056	0.057	0.068	0.094
40	0.088	0.099	0.095	0.096	0.089	0.093	0.159
70	0.177	0.176	0.131	0.137	0.189	0.126	0.202
Lower Foothills							
10	0.097	0.093	0.086	0.082	0.080	0.084	0.149
40	0.192	0.175	0.136	0.118	0.240	0.106	0.265
70	0.272	0.270	0.182	0.179	0.244	0.140	0.320
Montane							
10	0.056	0.069	0.085	0.088	0.073	0.089	0.059
40	0.064	0.073	0.131	0.130	0.074	0.140	0.123
70	0.095	0.093	0.147	0.147	0.088	0.152	0.170
Alpine							
10	0.061	0.069	0.057	0.060	0.074	0.072	0.068
40	0.137	0.130	0.111	0.102	0.120	0.111	0.158
70	0.166	0.150	0.116	0.118	0.161	0.112	0.205

Table D - 2: Summary statistics for the candidate noise reduction technique RMSE calculations.

	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	None
Minimum	0.035	0.050	0.050	0.052	0.049	0.057	0.027
Maximum	0.272	0.270	0.182	0.179	0.269	0.173	0.320
Mean	0.112	0.118	0.103	0.103	0.120	0.101	0.150
Standard Deviation	0.072	0.068	0.040	0.038	0.072	0.034	0.080

D.2. Metric Calculation Results

Table D - 3: Summary of metric calculation results, showing the mean, standard deviation and range of metrics derived from the original model NDVI time series and the seven candidate noise reduction strategies.

	Model	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Maximum NDVI								
Mean	0.762	0.691	0.643	0.718	0.731	0.645	0.760	0.769
Standard Deviation	0.069	0.117	0.120	0.093	0.097	0.151	0.100	0.091
Range	0.221	0.499	0.518	0.508	0.532	0.593	0.559	0.509
Time of Maximum NDVI								
Mean	13.1	13.5	13.6	13.9	14.3	13.4	14.0	13.3
Standard Deviation	0.9	1.5	1.5	1.5	1.5	1.4	1.3	2.1
Range	3.0	9.0	6.0	7.0	6.0	7.0	8.0	12.0
NDVI Amplitude								
Mean	0.646	0.629	0.560	0.615	0.629	0.566	0.687	0.743
Standard Deviation	0.082	0.113	0.109	0.094	0.096	0.143	0.103	0.097
Range	0.270	0.516	0.546	0.467	0.474	0.618	0.612	0.509
Time of SOS								
Mean	5.1	5.0	4.6	5.3	5.1	4.3	3.5	3.9
Standard Deviation	1.2	1.8	1.7	2.0	1.9	1.8	1.5	2.0
Range	4.0	7.0	8.0	8.0	8.0	8.0	7.0	10.0
Time of EOS								
Mean	18.9	19.6	20.4	19.8	19.9	20.5	21.1	18.6
Standard Deviation	0.3	1.2	1.0	1.1	1.1	1.2	0.5	1.8
Range	1.0	8.0	7.0	7.0	6.0	8.0	2.0	9.0

Continued on next page.

	Model	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Length of Growing Season								
Mean	14.8	15.6	16.8	15.5	15.8	17.2	18.5	15.8
Standard Deviation	1.2	2.1	1.9	2.4	2.3	2.1	1.4	2.9
Range	4.0	11.0	9.0	10.0	9.0	9.0	7.0	15.0
Maximum Green-Up								
Mean	0.215	0.135	0.112	0.154	0.141	0.104	0.123	0.393
Standard Deviation	0.056	0.043	0.035	0.087	0.070	0.026	0.028	0.170
Range	0.202	0.214	0.161	0.336	0.265	0.128	0.130	0.676
Time of Maximum Green-Up								
Mean	6.3	8.4	8.2	8.2	8.1	8.1	7.5	8.1
Standard Deviation	2.4	2.3	2.2	1.6	1.6	2.2	2.8	2.7
Range	6.0	9.0	9.0	7.0	6.0	8.0	10.0	10.0
I-NDVI								
Mean	125.8	112.6	111.4	126.6	127.7	113.7	135.1	112.8
Standard Deviation	16.5	30.1	30.2	29.2	29.9	33.5	29.3	28.9
Range	56.2	127.7	126.7	129.2	133.0	137.9	138.0	121.5
Average NDVI								
Mean	0.532	0.441	0.410	0.502	0.503	0.407	0.461	0.398
Standard Deviation	0.069	0.095	0.096	0.090	0.080	0.104	0.082	0.110
Range	0.250	0.377	0.411	0.434	0.410	0.414	0.391	0.455

APPENDIX E: UNSTANDARDIZED (RAW) PERFORMANCE SCORES

E.1. RMSE Performance Scores

Table E - 1: Unstandardized RMSE overall summary performance scores.

	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
RMSE	203	212	187	186	215	181	268

Table E - 2: Unstandardized RMSE performance scores, stratified by land cover.

Land Cover	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Grassland	42	47	33	33	47	32	53
Parkland	14	18	20	20	20	20	19
Boreal	32	35	28	30	34	29	45
Lower Foothills	56	54	41	38	56	40	66
Montane	22	23	36	37	23	38	35
Alpine	37	35	29	28	35	29	43

Table E - 3: Unstandardized RMSE performance scores, stratified by noise level.

Noise Level	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
10%	35	41	40	40	38	51	42
40%	66	68	64	62	74	60	96
70%	102	103	83	84	103	77	123

E.2. Metric Performance Scores

Table E - 4: Unstandardized metric overall summary performance scores.

	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Metrics	1246	1462	1258	1238	1495	1491	1666

Table E - 5: Unstandardized metric performance scores, stratified by land cover.

Land Cover	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Grassland	225	301	235	228	298	249	334
Parkland	191	219	170	180	211	219	173
Boreal	179	239	198	204	225	238	324
Lower Foothills	233	248	222	211	280	239	328
Montane	167	205	208	189	220	254	216
Alpine	251	250	225	226	261	292	291

Table E - 6: Unstandardized metric performance scores, stratified by noise level.

Noise Level	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
10%	286	357	335	316	360	448	318
40%	352	418	364	350	423	432	682
70%	455	513	389	407	535	466	666

Table E - 7: Unstandardized metric performance scores, stratified by year.

Year	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
2003	419	508	417	391	488	511	489
2004	376	440	392	390	501	475	555
2005	451	514	449	457	506	505	622

Table E - 8: Unstandardized metric performance scores, stratified by year.

Metric	4253H-Twice Filter	ARMD3-ARMA5 Filter	Asymmetric Gaussian Function	Double Logistic Function	MVI Filter	Savitky-Golay Filter	Noisy Data
Maximum NDVI	77	110	73	73	110	42	68
Timing of Maximum NDVI	49	59	81	89	52	56	74
NDVI Amplitude	92	113	79	76	105	121	96
Start of Season	153	174	170	165	177	204	209
End of Season	54	83	62	65	87	49	105
Length of Season	95	113	97	95	126	117	162
Maximum Green-Up	239	284	269	257	290	528	244
Timing of Maximum Green-Up	271	268	246	245	279	279	331
Integrated NDVI	98	110	95	95	116	106	99
Average NDVI	118	148	86	78	153	164	103