Retrieval of biophysical parameters from multi-sensoral remote sensing data, assimilated into the crop growth model CERES-Wheat

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der Mathematisch–Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von
Vanessa Heinzel
aus Witten–Herdecke

Bonn, Dezember 2007

Angefertigt mit Genehmigung der Mathematisch–Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

1. Gutachter: Prof. Dr. Gunter Menz
2. Gutachter: Prof. Dr. K. Greve
3. Gutachter: Prof. Dr. A. Skowronek
4. Gutachter: Prof. Dr. W. Kühbauch

Tag der mündlichen Prüfung:
08.04.2008

Contents

List of Figures

List of Tables

LIST OF ACRONYMS

AP	Alternating Polarization
ASAR	Advanced Synthetic Aperture Radar
ASTER	Advanced Spaceborne Thermal Emission And Reflection Radiometer
BACROS	Basic Crop Growth Simulator (model)
BBCH	Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie
BRDF	Bi-Directional Reflectance Factor
CAP	European Common Agriculture Policy
CCD	Charge Coupled Device
CERES	Crop Environment Resource Synthesis
CNES	Centre National d'Etudes Spatiales
CORINE	Coordination Of Information On The Environment
DLR	Deutsches Zentrum Für Luft- Und Raumfahrt
DM	Dry Matter Biomass
DN	Digital Numbers
DSSAT	Decision Support System For Agrotechnology Transfer
ELCROS	Elementary Crop Simulator (model)
Envisat	Environmental Satellite
ERS-2	European Remote Sensing Satellite 2
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation For Exploitation Of Meteorological Satellites
FM	Fresh Matter Biomass
FWHM	Full Width At Half Maximum
GIS	Geographic Information System
GMES	Global Monitoring For Environment And Security
GSM	Gravimetric Soil Moisture
GT	Ground Truth
HyMap	Hyperspectral Mapper
IR	Infrared Spectrum
LAI	Leaf Area Index

Landsat TM	Landsat Thematic Mapper
LUT	Look-Up-Table
MARS	Monitoring Agriculture By Remote Sensing Project
MIMICS	Michigan Microwave Canopy Scattering Model
NASA	National Aeronautics And Space Administration
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared Spectrum
OASIS	Optimizing Access To SPOT Infrastructure For Science
PAR	Photosynthetic Active Radiation
PPD	Plant Population Density
PRI	Precision Image
PROSPECT	Properties Spectra
PVI	Perpendicular Vegetation Index
PWC	Plant Water Content
QE	Quantum Efficiency
RADAR	Radio Detection And Ranging
RAR	Real Aperture Radar
RMSE	Root Mean Square Error
RSR	Relative Spectral Response
RT2	Second-Order Radiative Transfer Model
SAIL	Scattering Arbitrary Inclined Leaves
SAR	Synthetic Aperture Radar
SAVI	Soil Adjusted Vegetation Index
SEAMLESS	System For Environmental And Agricultural Model
SLR	Side-Looking Direction
SUCROS	Simple And Universal Crop Simulator (model)
SVAT	Soil-Vegetation-Atmosphere
SWIR	Shortwave Infrared
TIR	Thermal Infrared
VI	Vegetation Index
VIS	Visible Spectrum
VSM	Volumetric Soil Moisture
WCM	Water Cloud Model
WDVI	Weighted Difference Vegetation Index

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all the people, who have contributed to the realization of this thesis. You all have done this either full-consciously or unknowingly, but you can be sure that I appreciate every piece of your involvement.

The work was realized at the Center for Remote Sensing of Land Surfaces (ZFL) at the Rheinische-Friedrich-Wilhelms University of Bonn (Germany), within the framework of the ENVILAND project (FKZ 50EE0404) financed by the DLR/BMWI. Data were provided by an ESA CAT 1 proposal (C1P3115) and through an OASIS proposal (58-CE 6242). I would also like to express my gratitude to the DFG Research Training Group 722 “Information Technologies for Precision Plant Protection” for financial support.

First of all I would like to thank my supervisor Prof. Dr. G. Menz (Geography Institute, Univ. Bonn). He always showed great interest in my work and fully supported me.

I am very glad that I could win Prof. Dr. K. Greve (Geography Institute, Univ. Bonn) as my second supervisor. Thanks also goes to Prof. Skowronek and Prof. Kühbauch from the INRES Institute (Univ. Bonn).

Special thanks goes to Dr. M. Braun the administrator of the ZFL. He has risen my interest in the remote sensing topic. He has built up the ZFL and managed to create a friendly and open working atmosphere.

Björn Waske also belongs a special thanks for not only sharing the same office with me and being my colleague in the ENVILAND project, but also for having an open ear to all my questions. He was never tired of discussing whatever topic. His support accelerated my work progress tremendously.
Sebastian van der Linden (Schiefer) from the Humboldt University of Berlin (Germany) has also abetted me and was a great example for me.
I would also like to point out another important person, who was never tired off assisting me, Dr. J. Franke.
My other colleagues at the ZFL I would also like to thank for three years full of good discussions and a great working atmosphere: Dr. A. Moll, Dr. J. Jacobi, Roland Götzke, Ambros Waidosch and Michael Judex.
Ellen Götz helped me in all administrative questions and is actually the heart of the ZFL.

The staff at the research farm "‘Klein-Altendorf"’ I also owe many thanks, especially Mr. Huober. As well as all my field student assistance: Raimund, Jan, Thea, Claudia, Aline, Nina, Felix, Caroline, Susanne and all other student assistances.

Thanks also goes to the Pawlow, for the nice location, the good place for discussions and the tasty beer.

I would also like to thank Torsten Sander, who supported me during my Geography study and the first years of my PhD study. You will always be a very special person for me, I will never forget you.

Then I am indebted to my boyfriend Carl Jan Keuck. You gave me all the support and understanding I could wish for during the last critical months, both for my work and privately. You understood to cheer me up and to look forward to more relaxed times.

Finally, I would like to thank my parents Monika and Hans-Georg and my brother Sebastian for supporting and believing in me. From you I learned to stand up for my beliefs and feelings, no matter what others might say. You have risen my interest in nature science, creativity as well as social and political issues.

Once again, thank you all for your support and believe in me.

Bonn, December 2007

Vanessa Heinzel

ABSTRACT

This study investigated the possibilities and constraints for an integrated use of a crop growth model (CERES-Wheat) and earth observation techniques. The assimilation of information derived from earth observation sensors into crop growth models enables regional applications and may also help to improve the profound knowledge of the different involved processes and interactions. Both techniques can contribute to improved use of resources, reduced crop production risks, minimised environmental degradation, and increased farm income.

Up to now, crop growth modelling and remote sensing techniquices mostly have been used separately for the assessment of agricultural applications. Crop growth models have made valuable contributions to, e.g., yield forecasting or to management decision support systems. Likewise, remote sensing techniques were successfully utilized in classification of agricultural areas or in the quantification of vegetation characteristics at various spatial and temporal scales. Multisensoral remote sensing approaches for the quantification biophysical variables are rarely realized. Normally the fusion of the data sources is based on the use of one sensor for classification purposes and the other one for the extraction of the desired parameters, based on the map classified previously. Pixel-based fusions between multispectral and SAR data is seldom realised for the assessment of quantitative parameters.

The integration of crop growth models and remote sensing techniques by assimilating remotely sensed parameters into the models, is also still an issue of research. Especially, the integration of, e.g., multi-sensor biophysical parameter time-series for the improvement of the model performance, might feature a high potential.

The starting point of the presented study was the question, if it is possible to derive the values of important crop variables from various remote sensing data? For the retrieval of these quantitative parameters by the use of various multispectral remote sensing sensors, intercalibration issues between the different retrieved vegetation indices had to be taken into account, in order to assure the comparability. Features influencing the vegetation indices are, e.g., the sensor geometry (like viewing- and solar-angle), atmospherical conditions, topography and spatial or radiometric resolution.

However, the factors taken into account within this study are the spectral characteristics of the different sensors, like band position, bandwidth and centre wavelengths, which are described by the relative spectral response functions. Due to different RSR functions of the sensor bands, measured spectral differences occur, because the sensors record different components of the reflectance's spectra from the monitored targets. These are then also introduced into the derived vegetation indices. The chosen cross-calibration method, intercalibrated the assessed Normalized Difference Vegetation Index and the Weighted Difference Vegetation Index between the various sensor pairs by regression, based on simulated multispectral sensors. Differences between the various assessed remote sensing sensors decreased form around 7% to below 1%. The intercalibration also had a positive impact on the later biophysical retrieval performance, producing sounder retrieval results.

For the retrieval of the biophysical parameters empirical and semi-empirical models were as-

sessed. The results indicate that the semi-empirical CLAIR model outperforms the empirical approaches. Not only for the Leaf Area Index retrieval, but also in the cases of all other assessed parameters.
Concerning the other remote sensing data type used, the SAR data, it was analysed what potential different polarizations and incidence angles have for the extraction of the quantitative parameters. It became obvious that especially high incidence angles, as provided by the satellite Envisat ASAR, produce sounder retrieval results than lower incidence angles, due to a smaller amount of received soil signal. In the context of the assessed polarizations, sound results for the VV polarization could only be achieved for the retrieval of fresh biomass and the plant water content. For the ASAR sensor modelling fresh biomass and LAI using the HV polarization or the dry biomass using the ratio (HH/HV) was appropriate.
As roughness aspects also have an influence on the retrieval performance from biophysical parameters using SAR data, the impact of soil surface and vegetation roughness was additionally considered. Best results were achieved, when also considering roughness features, however due to the need of regional modelling it is more appropriate not to consider them.
For the calibration and re-tuning of crop growth models information about important phenological events such as heading/flowering is rather important. After this stage reproductive growth begins, whereby the number of kernels per plant is often calculated from plant weight at flowering and kernel weight is calculated from time and temperature available for dry matter distribution. By the use of the SAR VV time-series this important stage could be successfully extracted.
Further methods for pixel-based fused biophysical parameter estimations, using SAR and multispectral data were analysed. By this approach the different features, being monitored of the two systems, are combined for sounder parameter retrieval. The assessed method of combining the multi-sensoral information by linear regression did not bring sound results and was outperformed by single sensor use, only taking into account the multispectral information. Only for the parameter fresh biomass, modelling based on the NDIV and the ASAR ratio slightly outperformed the single sensor modelling approaches. The complex combined modelling by the use of the CLAIR and the Water Cloud Model featured no valid results. For the combination, by using the CLAIR model and multiple regression slight improvements, in contrast to the single multispectral sensor use, were achieved. Especially, during late phenological stages, the assessed VV information improved the modelling results, in comparison to only using the CLAIR model.
All the findings could finally be successfully applied for regional estimations. Only the roughness features could not be applied, due to the fact, that it is hard to regionally assess this needed model input parameter. Regional parameter on the basis of remote sensing data, is the major advantage of this technique, due to the large spatial overview given.

The second main question was, if it is possible to integrate the crop variables gained from multisensoral data into a crop growth model, increasing the final yield estimation accuracy. Thus far, beneficial linkages between both techniques have been often limited to land use classification via remote sensing for choosing the adequate model and quantification of crop growth and development curves using biophysical parameters derived from remote sensing images for model calibration. Only a few studies actually considered the potentials of remote sensing for model re-initialization of growth and development characteristics of a specific crop, as the here studied winter wheat. Overall, the integration of remotely sensed variables into the crop growth model CERES-Wheat led to an improved final yield estimation accuracy in comparison to an automatic input parameter setting. The assessed final yield bias for the automatic input parameter setting summed up to 6.6%. When re-initializing the most sensitive input parameters (sowing date and

fertilizer application date) by the use of remotely sensed biophysical variables the biases ranged from 0.56% overestimation to 5.4% understimation, in dependence of the data series used for assimilation. Whereby, it was assessed that the combined dense data series, considering SAR and multispectral information, slightly outperformed the performance of the full multispectral data series. However, when analysing the assimilation of the multispectral data series in further detail, it became clear that the actually information from the phenological stage ripening declines the modelling performance and thus the final yield estimation accuracy. When neglecting the information from this phenological stage the reduced multispectral data series performed as sound as the dense data series containing SAR and multispectral information. Thus, when the appropriate phenological stages are monitored by multispectral data, additional SAR information does not lead to a model improvement. However, when important dates are not monitored by multispectral images, e.g., due to cloud coverage, the additionally considered SAR information was not able to appropriatly fill these important multispectral time gaps. They even had a more negeative influence on the modelling performance. Overall, the best results could be obtained by assimilating a multispectral data series, covering the crop development during the important phenological stages stem elongation and flowering (without ripening stage), into the CERES-Wheat model.
Finally, the integration of remote sensing data in the point-based crop growth model allowed it‘s spatial application for prediction of wheat production at a more regional scale. This approach also outperformed another evaluated method of direct multi-sensoral regional yield estimation.

This study has demonstrated that biophysical parameters can be retrieved from remote sensing data and led, when assimilated into a crop growth model, to an improved final yield estimation. However, overall the SAR information did not really have a significant positive effect on the multi-sensoral biophysical parameter retrieval and on the later assimilation process. Thus, overall SAR information should only be considered, when multispectral data acquisitions are tremendously hampered by cloud coverage. The assessed assimilation of remote sensing information into a crop growth model had a positive effect on the final yield estimation performance. The analysed method, combining remote sensing and crop growth model techniques, was successfully demonstrated and will gain even more importance in the future for, e.g., decision support systems fine-tuning fertilizer regimes and thus contributing to more environmentally sound and sustained agricultural production.

1 INTRODUCTION

1.1 Background

Land-use activities, whether converting natural landscapes for human use or changing management practices on human-dominated lands, have transformed a large proportion of the planet´s land surface (Foley et al., 2005). Crop lands and pastures have become one of the largest terrestrial biomes on the planet occupying around 40% of the land surface. Modern agriculture now feeds six billion people, whereby global cereal production has doubled within the past 40 years (Tilman et al., 2002). Some of this increase can be attributed to a twelve % increase in world crop land, but it is mainly due to increased yields resulting from greater inputs of fertilizer, water and pesticides, new crop strains, and other technologies of the *Green Revolution*. Within this time period there has been a 700% increase in fertilizer use and a 70% increase in irrigated crop land area. This has enlarged the global food supply, reducing hunger, improving nutrition and thus the ability of people to better reach their mental and physical potential, and sparing natural ecosystems from conversion to agriculture (Ciais et al., 2005; Foley et al., 2005).
Even so modern agriculture has been successful in increasing food production, it has also caused extensive environmental damage, e.g., degradation of water quality, heavily salinized arable land, soil erosion or loss of native habitats (Wood et al., 2000). Modern agricultural land-use practices may be trading short-term increase in food production for long-term losses in ecosystem services.
By 2050, the global population is projected to be 50% larger than at present and global grain demand will also double (Cassman, 1999; Tilman et al., 2001; Waggoner, 1995). Further increases in agricultural output are essential for global political and social stability and equity. Doubling food production again and keeping food production at this level, are major challenges (Tilman et al., 2002). In this context there are a number of different national and international programs, aiming at the enhancement of food security.
The member states of the United Nations committed themselves in 2000 to eight quantifiable Millennium Development Goals (MDGs), until 2015, in order to improve the livelihoods for millions of people (Worldbank, 2000). 70% of the MDG´s target group live in rural areas, thus welfare may be achieved through agriculture. The linkage with agriculture is the strongest for the first MDG (eradicating poverty and hunger), however all other MDGs also have direct or indirect linkages with agriculture. In MDG7 (ensure environmental sustainability) agricultural practices can be both direct causes of, and important solutions to environmental degradation. This is also valid for poor rural areas in less developed countries, but also for agriculture in developed countries.
The EU/ESA Global Monitoring for Environment and Security program (GMES)[1] aims at the establishment of an operational service for crop monitoring in support of food security monitor-

[1] http://www.gmes.info

ing to serve policy makers and operational users. The project has hereby an inclination towards the use of Earth observation data where appropriate. The services provided cover three major categories:

- Mapping, including topography or road maps, but also land-use and harvest, forestry monitoring, mineral and water resources that do contribute to short and long-term management of territories and natural resources
- Support for emergency management in case of natural hazards and particularly civil protection institutions responsible for the security of people and property
- Forecasting, applied for marine zones, air quality or crop yield

In 1988 the Council of Ministers of the European Union (EU) set up a project to improve the provision of agricultural statistics, which are necessary to manage the large budgets involved in the European Common Agricultural Policy (CAP). This project is known as the Monitoring Agriculture by Remote Sensing (MARS) and compromises different activities such as regional crop inventories, satellite-based rapid crop area estimates, assessment of foreign agricultural production and an agricultural information system[2]. The agricultural information system activity focuses on providing early crop yield forecasts for the EU countries by either using remote sensing data or an agrometeorologic system employing crop growth models.
Also a recently started project is the SEAMLESS project (System for Environmental and Agricultural Modelling; Linking European Science and Society)[3]. It aims at analysing the full range of scale (farm to EU and global), while focusing on the most important issues emerging at each scale. One other important issue is the analysis of the environmental, economic and social contributions of a multifunctional agriculture towards sustainable rural development and rural viability. The analysis of a broad range of issues and agents of change, such as climate change, is also assessed in the project.
Remote sensing plays an important role within this context, providing temporal and spatial information. By coupling the acquired information with, e.g., crop growth models, it can assist in assessing variability in crop performance and final yield and can even provide information of, and for, e.g., agricultural management interference.

1.2 Paradigm change in the remote sensing community

Remote sensing as a technology can actually be dated to the middle of the 19th century, with the first aerial photographs taken from a balloon over Paris. Photography has served as a prime remote sensor for more than 150 years. In the 1930s RADAR technologies were developed and in the 1950s remote sensing systems continued to evolve from the systems developed for warfare (Jensen, 2007).
The first sophisticated imaging sensors were incorporated in orbiting satellites in the course of the 1960s. Firstly, only being basic television cameras, imaging low-resolution black and white

[2] http://www.ena.lu/
[3] http://www.seamless.eu

pictures of clouds and the earth´s surface. Since then, other sensor types were developed, also taking images using the electromagnetic spectrum beyond the visible.
A big variety of different systems monitoring the land surface in different regions of the wavelength spectrum with different spatial coverage and resolution have been developed, like the Landsat missions or the commercial satellites Ikonos and QuickBird.
The successes over the last two decades and the most prominent European civilian achievements include[4]: The meteorological satellite systems operated by EUMETSAT (European Organisation for Exploitation of Meteorological Satellites). They deliver weather and climate-related satellite data and products, 24 hours a day all year long. The ESA Envisat satellite, using advanced all-weather radar technologies and multispectral observation techniques in order to provide continuous observation and monitoring of the Earth´s land, atmosphere, oceans and ice caps. The ESA ERS satellites carrying a comprehensive payload including an imaging Synthetic Aperture Radar, a radar altimeter and other powerful instruments to measure ocean surface temperature and winds at sea. The recently launched TerraSAR-X radar satellite, the first German satellite realized in Public Private Partnership (German Ministry of Education and Science, German Aerospace Center, Astrium GmbH). The mission´s objectives are the provision of data for scientific research, applications, the establishment of a commercial market and the development of a sustainable EO-service business. The German satellite constellation RapidEye (5 identical satellites), will be launched in summer 2008. The constellation will be the only commercial Earth observing system capable of providing large-area multispectral coverage at high resolution on a daily basis. The application services range from disciplines as Agriculture to Forestry, Power and Communications, Spatial Solutions or Government. The last prominent achievement in Europe is the Sentinel Family developed by the ESA. It will consist of five different series (SAR, Superspectral, Ocean, 2 time Atmospheric Chemistry), whereby the operation concept is based on the requirements for an operational GMES system. There are of course many other satellite missions outside of Europe, but summarizing them all would go beyond the scope of this chapter.
Overall one can state that there rapid increase in:

- number of missions and constellations
- number of sensors
- kinds of sensed data
- sensor resolution
- number of spectral bands
- number of data formats
- number, type and size of distributed archives
- and additionally available information in geographical information systems

Nowadays, there is a variety of different remote sensing sensors in the orbit, acquiring data in different regions of the electromagnetic spectrum, with different spatial resolution and coverage

[4] http://www.gmes.info

and due to pivoted sensors with the possibility of imaging respective regions of the land surface in higher frequencies. Due to this increased potential, there is a slight paradigm change in the remote sensing community. Information mining/knowledge discovery and the associated data management are changing the paradigms of the user/data interactions. New technologies and work flows are required to automatically analyse such data and data series (Dactu et al., 2002). Generally, we are now at the stage, where we have a range of successful and widely used analytical procedures and an abundance of data. This is quite contrary to the situation at the start of the space-sensing era. Today the analyst is challenged to choose from among a number of coincident data sets when undertaking, e.g., thematic mapping (Richards, 2005).
While the analysis used to be sensor driven, meaning that the available sensors dictated the products and assessed questions, today the research questions and the desired products dictate the remote sensors to be used. In a modern operational setting, the application requirements would be specified by a client, and a consultant would have the task of choosing the datasets and analytical methods for generating the desired product. Remote sensing, fundamentally, is an applications driven field, and while there is still room for the development of further analysis algorithms, the requirements of the end user will drive the outcomes. The methodologies will now be as much about choosing the most relevant primary data types, as well as performing the actual analysis (Richards, 2005).
Within this context also the fusion of the different data sets has become an important aspect, whereby the advantages of each sensor system, e.g., microwave and multispectal systems, are used for the improvement of the accuracy of the derived information. Image fusion approaches can be divided into three main categories, based on the stage at which the fusion is performed: pixel-based, feature-based and decision-based (Pohl & van Genderen, 1998).
In pixel-based fusion, the data are merged on a pixel-by-pixel basis. This can be achieved by using both data types in the same algorithm or by combining, e.g., the time-series of the data, enabling to better monitor dynamic processes like crop growth or for classification purposes (Brakke et al., 1981; Liu et al., 2006). A last rather common option is, e.g., really combining the different data types on the pixel basis using different algorithms (e.g., wavelet fusion, IHS-transformation). Feature-based approaches merge the different data sources at the intermediate level. It requires the extraction of objects recognised in the various data sources, e.g., using segmentation procedures. The last combination option is the decision-based fusion, the outputs of each of the single source interpretations are combined to create a new interpretation, e.g., the combination of different classification results on the basis of the class labels (Waske & Benediktsson, 2007). Within the presented study the fusion was mainly done pixel based, thus at an early analysis stage.

1.3 Challenges of combining models and remote sensing information

Information technologies as simulation modelling and remote sensing play an important role in agricultural research and in agricultural production systems. A substantial technical progress has been made in both, which is still continuing, due to the increasing possibilities in computer hard- and software and technological progress, combined with greater knowledge on modelled and observed systems (Jongschaap, 2006). Although remote sensing and crop growth modelling have proven their usefulness and applicability, such approaches are seldom combined.

Concerning the remote sensing issue, the retrieval of bio- and geophysical parameters is an important field of research, and the prospect of extracting such information in an operational manner with a high degree of accuracy has a strong impact on current scientific work (Rosenqvist et al., 2003). Meaningful parameter retrieval requires not only the availability of appropriate inversion algorithms, but also that locally developed models can be applied to an extensive regional context. The issues of operational use and regional transfer (extraction to larger regional areas) are important, when wanting to combine remotely sensed information with crop growth models for regional analysis.
Increased process knowledge and the concurrent improvements have led to the development of complicated crop growth models, needing a rather large amount of input parameters. This makes it rather hard to apply the point-based model findings to a regional spatial context. Here remotely-sensed information can be linked successfully to the crop models and help to overcome this problem.
Combinations of remote sensing and simulation modelling can be synergetic in various ways. The areas of interest within this thesis were:

- the use of estimates of biophysical variables that can be used to re-initialize the crop model

- using the spatial aspect of remote sensing images to upscale the simulation results to obtain regional results

1.4 Geographic context of the thesis

The main attention of the presented thesis is the contention with the retrieval of biophysical parameters using different remote sensing data. The role of remote sensing data within the field of geography experienced an increasing importance. It lies in the nature of the issue that Geography, dealing with manifold issues of the land surface, can gain spatial information about the earth surface and it‘s interactions by interpreting and analysing remote sensing data (Albertz, 2001). The part of remote sensing is quite diverse and also connected with lots of other research disciplines as Geology or Cartography. The principle geographic method of aerial or remote sensing analysis dates back to Carl Troll, during the second world war. It made advantage of a large regional overview of data, to analysing different landscape elements. The aim was to detect causes and effects of landscape elements, in order to identify functional relationships within the landscape.
Nowadays, the possibilities of remote sensing data have become even more important and new dimensions, through acquiring data in different wavelength regions or multi-temporal data, also with higher spatial resolution.
Remote sensing is a tool for mapping the world and thus detecting either land cover changes or for assessing the actual land use. This is not only important for process modelling, but also for understanding, e.g., impacts of political decisions. Also in cases of disaster monitoring, e.g., after the tsunami in December 2004 remote sensing played a dominant role for not only assessing the effects of the disaster, but also for the organization of first aid help. Another important topic within the Geographic content is climate change, whereby remote sensing data are used for monitoring large areas and to detect changes over time (Jensen, 2007).
The second focus of this thesis is the assimilation of the remotely derived information into crop

growth models for final yield estimation. The context of modelling in order to understand processes has a long tradition in the geographic community, whereby a great variety of models are used, e.g., hydrological models for estimating surface run off, or climate models for assessing the impact of, e.g., CO_2 in the atmosphere.
Agrargeographic hypotheses generally have a long tradition within the geography, e.g., the well known *Thünschen Ringe* (von Thünen, 1910). Especially, since the structural changes within the primary sector the interest of geographers in agricultural issues has increased. The Agrargeography deals with the description and analysis of agricultural dominated areas, their materials, shape, geographic position, structure as well as their process relationships and strengths (Arnold, 1997). It analyses the interactions between natural (climate, relief, soil, water) and anthropogen geofactors (politics, war, economics, population, settlements). Whereby, nowadays also regional yield estimations play an important role not only for improved disaster management, or within the European report duties, but also for economic planning.
Overall, the thesis has many connections to Geography, whereby probably the strongest one is the handling of spatial information for modelling interactions in the earth environment.

1.5 Objectives and hypotheses

There is a need for the development of methodologies with respect to the integration of soil and crop processes, expert knowledge and observations from remote sensing techniques, to more adequately understand the crop performance and for better agricultural management. Many new remote sensing sensors and techniques have been developed in recent years for collecting information on crop and soil systems in various spectral bands, hyperspectral, laser-induced fluorescencs, microwave, etc.. Due to the fact that remotely sensed information is only a representative of the actual soil-crop status, it is difficult to identify the processes responsible for the observed crop and soil conditions. However, the identification of these processes is essential for the selection of the appropriate intervention through management. In the context of such selections mono-temporal remote sensing observations are hard to interpret without using additional information on crop and soil status. Multi-temporal measurements improve the possibilities for identifying the relevant processes, if reliable methods are available for processing and interpreting the data. In order to maximize the possibilities of identifying the appropriate management process simulation models can be linked to the remotely sensed information. The success of doing this largely depends on the accuracy of the retrieved remotely sensed information. Thus, different retrieval techniques for multispectral and SAR data, as well as the combination of these sensors for biophysical information extraction and their implementation into a process model are explored and discussed in this thesis. Although remote sensing and crop growth modelling both have proven their usefulness and applicability in various areas, their use is seldom combined. The challenges are to analyse the effects of integrating a time-series of multispectral and microwave information into crop growth models. The methods of biophysical parameter retrieval are plentiful but still, especially when combining multi-sensoral data, not very well assessed. The used remote sensing models have to stay rather simple for inversion purposes and the ability of applying them to larger regions. The models have to be operationally and easy, based on simply to obtain parameters. Also the impact on crop growth modelling when integrating the retrieved information is still not clearly analysed. With regard to the current scientific status the first challenging question within this thesis was:

Is it possible to derive the values of important crop variables from various remote sensing data?

Whereby, in more detail the following aims were of concern:

- to intercalibrate vegetation indices of different multispectral sensors in order to correct the biases, introduced by differences in the relative spectral response functions (section 4.1.1)
- to evaluate the performance of the CLAIR model (Clevers, 1989) for the retrieval of different biophysical parameters (section 4.1.3)
- to analyse the potential of different polarizations and incidence angles for the retrieval of different biophysical parameters (section 4.2.1)
- to analyse the performance of the Water Cloud Model (Attema & Ulaby, 1978) and multiple regressions using different SAR data and the additional variable *vegetation roughness* (section 4.2.2 and 4.2.3)
- to extract important phenological events such as heading/flowering from remote sensing information for the adjustment of the retrieval models and the simulation model (section 4.2.4)
- to analyse the potential of combining multispectral and SAR data for the derivation of biophysical parameters (section 4.3)
- to invert the used models for regional biophysical parameter extraction

These derived biophysical parameters were then assimilated into a crop growth model for final yield estimation, whereby the question was:

Is it possible to integrate important crop variables gained from multi-sensoral data into a dynamic crop growth model in order to increase the final yield estimation accuracy?

Challenges hereby were:

- to assimilate biophysical parameter maps retrieved from different remote sensing data into a crop growth model in order to improve the accuracy of wheat yield predictions (section 5.4)
- to analyse the effect of assimilating separately or synergistically SAR and multispectral time-series information (section 5.4)
- to analyse the effect of acquisition time and frequency of the assimilated multispectral data (section 5.5)
- to assess whether time gaps in the multispectral data series can be filled by SAR information (section 5.6)

- to transfer the findings of the point-based crop growth model for regional final yield prediction (section 5.8)

- to analyse the performance of the suggested approach in comparison to a simple direct yield estimation method (section 5.9)

1.6 Synopsis

In Fig. 1.1 the organisation structure of the thesis is shown. Previous to all analysis steps and all further considerations in chapter **THEORY AND RESPECTIVE STATUS OF SCIENCE** (2) the necessary theoretical background will be given. Whereby, first considerations about multispectral remote sensing and afterwards about SAR remote sensing will be made. Concerning multispectral remote sensing only some important aspects, as multispectral features of vegetation and multi-temporal signature characteristics, as well as differences in relative spectral response functions from the various used multispectral sensors, will be highlighted. This is followed by an introduction to the topic SAR remote sensing, describing special features of vegetation and soil interactions with the backscatter and multi-temporal aspects, which have to be taken into consideration. The next aspect of the chapter will deal with the retrieval of biophysical parameters from the various remote sensing data, whereby also statistical features will be summarized. These biophysical parameters will then be assimilated into the crop growth model CERES-Wheat. For the assimilation process, different techniques and the CERES-Wheat model will be introduced, as well as some features of the agricultural plant winter wheat. Thus, overall there will be three different important theoretical sections. One highlighting the sensor/target interactions, the second one discussing methods for the derivation of biophysical parameters from the different remote sensing sensors and the third one dealing with crop growth simulation modelling and methods for assimilating remotely sensed information into these models.
After the introduction of the theoretical background the different data sets will be described, which will be the collected ground truth data, as well as needed additional data and the different used remote sensing data (chapter **DATA** 3). By the use of these data in the chapter **BIOPHYSICAL PARAMETER ESTIMATION** (4) the biophysical parameters of interest will be derived from the multispectral time-series, by using empirical regressions and the semi-empirical CLAIR model (section 4.1). Therefore, an analysis regarding the intercalibration of the different used vegetation indices will be made, taking into account differences caused by variations in the relative spectral response functions of the respective multispectral remote sensors. In the next sections 4.2.2 and 4.2.3 the biophysical parameters will be extracted from the SAR time-series and the retrieval performances will be discussed. In the section 4.3 the possibilities of a combined retrieval approach will be analysed using the multispectral and SAR data in a synergetic way. In the chapter **CERES-WHEAT CROP GROWTH MODELLING AND FINAL YIELD ESTIMATION** (5) the remotely sensed information will be assimilated into the CERES-Wheat model. Whereby, the modelling performances of the different assimilated data (section 5.4) and the frequency and acquisition time of the multispectral data will also be considered (section 5.5). Another important aspect, which will be highlighted is, if SAR information is able to fill the time gaps in the multispectral data series, from distinctive phenological stages (section 5.6). The findings will then be used for a regional final yield estimation in the subsection 5.8 and finally they will be compared with a method of direct final yield estimation in subsection 5.9.

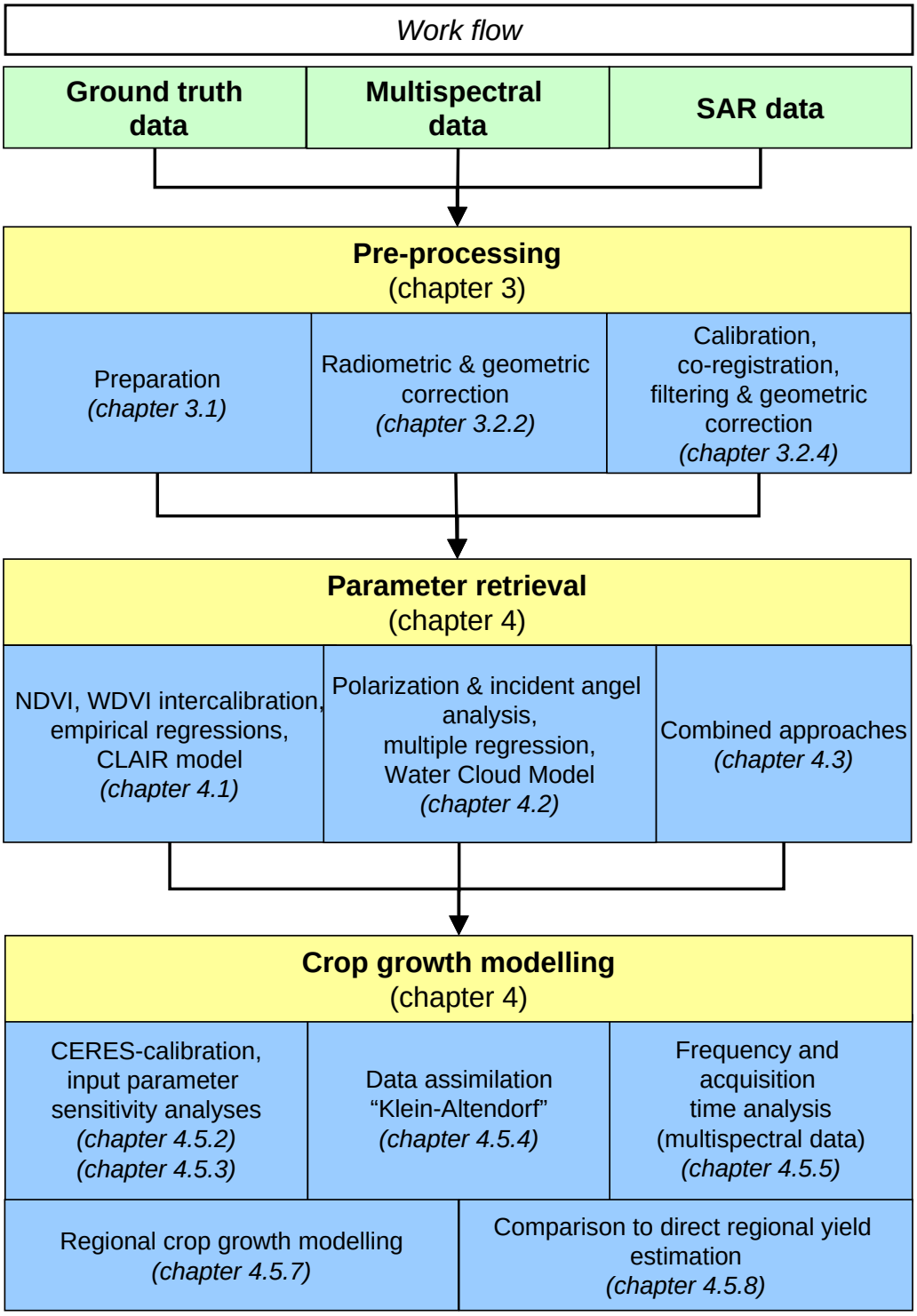

Figure 1.1: Graphical overview of the thesis organization.

2 THEORY AND RESPECTIVE STATUS OF SCIENCE

In this chapter the necessary theoretical context and literature overview will be given. In the first section considerations about remote sensing (multispectral and SAR features) and agricultural vegetation will be made. In the next section the retrieval of important biophysical parameters from these different data types will be explained. The retrieved parameters will be the inputs for the crop growth modelling. Deliberations about the assimilation of remotely sensed parameters into crop growth models will be made. After this theoretical introduction, the thesis will be arranged into the context of the current scientific status and the challenges within the study will be described.

2.1 Remote sensing and agricultural vegetation

2.1.1 Multispectral characteristics of vegetation

Multispectral remote sensing has been a research topic for many years, thus it is documented in many textbooks. Due to this fact, the introduction will be rather short only highlighting some features, which were of concern within the study.
The interested reader is referred to, e.g., Lillesand & Kiefer (2000); Mather (1999); Schowengerdt (1997); Ustin (2004) for a detailed introduction of multispectral remote sensing.

2.1.1.1 Multispectral features of vegetation

Especially important for vegetation monitoring and the retrieval of vegetation characteristics is the reflectance difference between the red and the infrared spectral wavelengths of leaves. The electromagnetic radiation gets reflected, transmitted and absorbed by healthy leaves in dependency of the wavelength and leave characteristics, e.g., morphology, anatomy, physiological features. Typical physiological features are leave structure, organization of the cells and pigments.
Vital leaves feature a typical distribution of reflection, transmission and absorption. In general the typical spectral reflectance curves of healthy leaves can be ascribed to different processes within the plant. Pigments in the plant leaves (chlorophyll) strongly absorb energy in the wavelengths at $0.45\mu m$ and $0.67\mu m$. If a plant is exposed to stress interrupting it‘s normal growth it decreases chlorophyll production and therefore the absorption in the blue and green wavelength

spectrum. Between $0.7\mu m$ and $1.3\mu m$ the reflectance of healthy vegetation usually ranges from 40 to 50%, due to the internal structure (mesophyl) of the plants. The dips in the spectral curves beyond $1.3\mu m$ can be ascriped to the water content of the leaves.
The Fig. 2.1 displays a spectral vegetation profile taken from a hyperspectral HYMAP imagery together with the mentioned effects (Lillesand & Kiefer, 2000). Important for the interpretation of the remote sensing data, is that the reflectance characteristic of a single leave is not the same as for the whole crop canopy. Due to the 3-dimensional structure of crops features like, e.g., leaf inclination, shadow, multiple leaf or soil scattering, also have an impact on the reflectance signal (Bauer et al., 1986). These features are especially important when observing, e.g., whole agricultural fields.

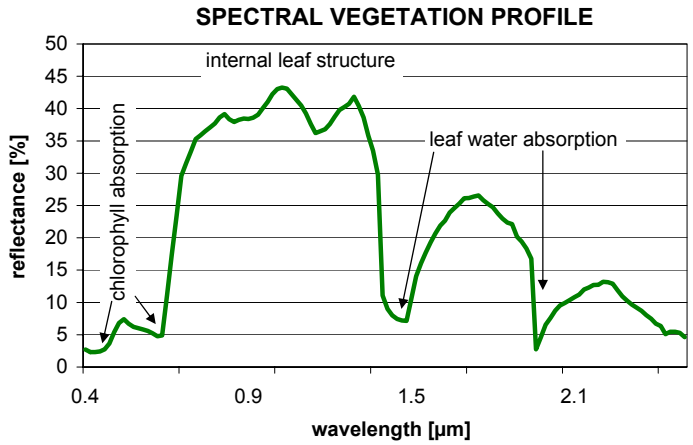

Figure 2.1: Typical spectral vegetation profile, measured by the hyperspectral HYMAP sensor (05/28/2005), modified after: Lillesand & Kiefer (2000).

2.1.1.2 Multi-temporal winter wheat reflectance profile

The crop phenological development and status can be monitored through multi-temporal reflectance profiles or multi-temporal vegetation indices (VI), e.g., the Normalized Difference Vegetation Index (NDVI). In Fig. 2.2 a typical seasonal NDVI profile of winter wheat is shown (Fang et al., 2005). The start of the season (SOS1) begins around October after planting. Due to the transition from bare soil to soil with varying amounts of canopy, the red and NIR patterns as well as the VI patterns change. The second start of the season (SOS2) begins in February, after the winter break (EOS1). There is typically a small VI decrease during the winter break. The end of growing season (EOS2) is marked trough minimum NDVI values. The growing season length (GSL) is defined as the distance between SOS and EOS. The beginning of the season is generally associated with rather low VI values. Within the crop development the values increase achieving a peak with the maximum crop development stage. After that, the fresh biomass decreases due to crop maturity and thus the VI values decrease. Due to different climate, soils or agricultural practices the crop development varies and thus the VI curves (Fontana et al., 2005; Ustin, 2004).

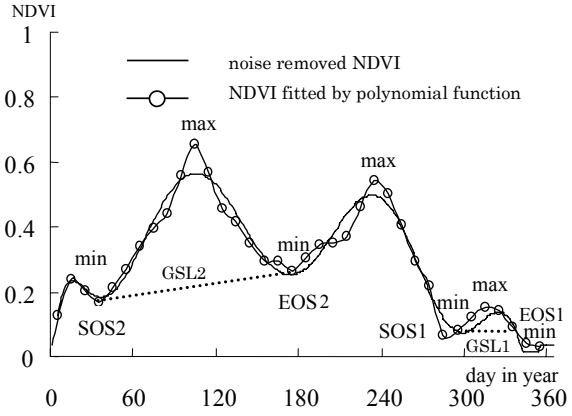

Figure 2.2: Seasonal NDVI profile of winter wheat (Fang et al., 2005)
(SOS: start of the growing season; GSL: growing season length; EOS: end of growing season)

2.1.1.3 Spectral sensor characteristics influencing the vegetation signal

For vegetation analyses often various sensors of different platforms have to be used, due to the low temporal resolution of satellites, e.g., 16 days repetition for Landsat 5TM and the problem of cloud coverage. As a consequence only the utilization of multiple platforms during a growing season enables at least a chance for several cloud free acquisitions for crop monitoring.
These multi-sensoral observations require considerations of some sensor characteristics influencing the spectral vegetation profile, in order to assure the comparability. For example, the sensor geometry, like viewing- and solar-angle, atmospherical conditions, topography and spatial or radiometric resolution influence the observed signal.
Other important factors are the spectral characteristics of the different sensors, like band position, bandwidth and centre wavelengths, which are described by the relative spectral response (RSR) functions. Due to different RSR functions of the sensor bands, measured spectral differences occur, because the sensors record different components of the reflectance's spectra from the monitored targets.
In general a spectral band of a sensor can be described by it´s spectral range, bandwidth, centre wavelength and full width at half maximum (FWHM). More specific it can be characterized by its RSR function, which takes into account all features describing a spectral band. The RSR function is affected by the effective spectral quantum efficiency (QE) of the detector. It includes not only the type-dependent sensitivity of the charge coupled device (CCD), but also losses due to the light reflecting or transmitting components of the detector (e.g., optics, mirrors, filters, etc.). Thus, variable sensor systems have different spectral sensitivity, which are described by their individual RSR functions (Franke et al., 2006).
Fig. 2.3 displays the RSR functions of the different sensors used in this study. Typical spectra of agricultural sites are plotted as reference. The functions differ in shape, central wavelength location and the degree of overlap between the VIS and NIR band. Especially, in the region of the red edge transition (0.68 to 0.8μm) the sensors differ from each other. Hence, this has an impact on the measured signal and the analysis of the observed spectral signal and has to be taken into account when comparing between different multispectral systems (Franke et al., 2006; van Leeuwen et al., 2006; Miura & Yoshioka, 2006; Steven et al., 2003; Trishchenko et al., 2002).

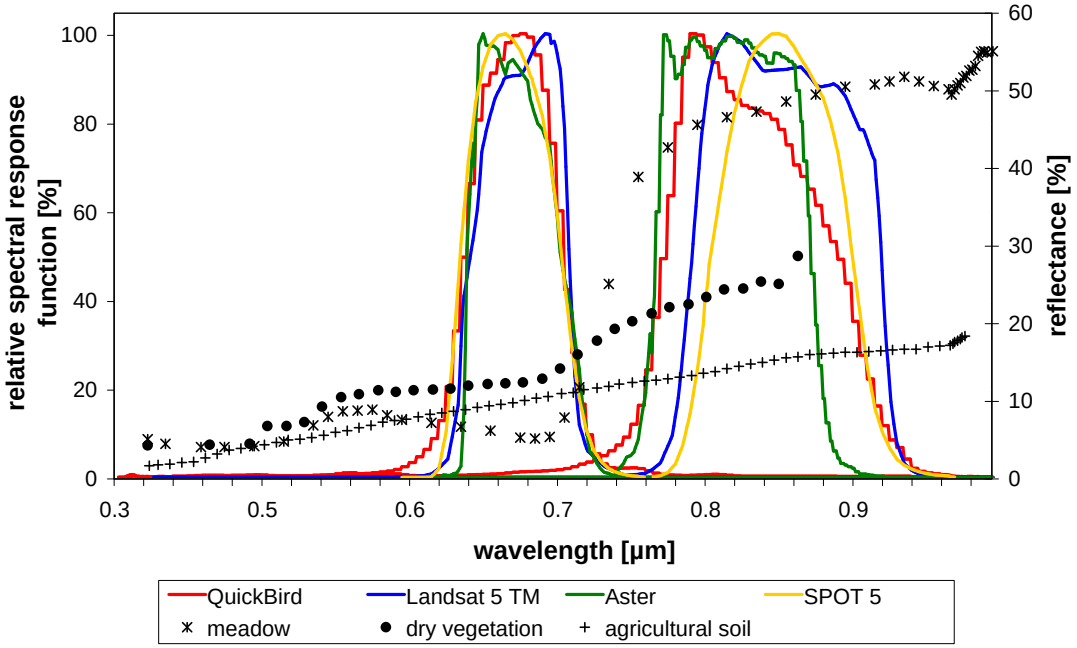

Figure 2.3: Variations in the relative spectral response curves of the satellites used within the study and spectral reflectance curves of some reference targets.

2.1.2 Microwave remote sensing fundamentals

Microwave remote sensing is fundamentally different from multispectral remote sensing, regarding, e.g., sensor characteristics, target characteristics and thus the information content in the received backscattered signal (Brisco & Brown, 1998; Oliver & Quegan, 2004; Ulaby et al., 1982; Woodhouse, 2006). By using the microwave region of the electromagnetic spectrum, capabilities are gained that complement other remote sensing methods, because interactions are driven by different physical interactions. E.g., the amount of microwave energy of a particular wavelength scattered off a green leaf is proportional to it´s size, shape and water content, rather than the amount of chlorophyll. Tab. 2.1 gives a brief overview of the different characteristics of multispectral and SAR waves. A detailed introduction into microwave remote sensing will be given in the following sections.

RADAR is an acronym for *Radio Detection and Ranging*, an active device transmitting and receiving electromagnetic energy in microwave wavelengths between 1mm and 1m. Typical characteristics of these systems are the ability to penetrate clouds, getting through the top layer of dry soils and as you provide your own illumination you are independent from the sun. Disadvantages on the other side are: the large required antennas for recording long wavelength and *S*ynthetic *A*perture *R*adar instruments tend to be the heaviest, largest, most power consuming and the data interpretation is mostly not that straight forward.

Table 2.1: Remote sensing with multispectral (VIS and IR) and SAR; crop characteristics influencing spectral reflection and microwaves (Kühbauch & Hawlitschka, 2003).

VIS and IR (λ, θ, ϕ)	SAR $(\lambda, \theta, \rho, \phi)$
pigment composition	volume (stand height)
pigment concentration	vertical and horizontal distribution of plant organs
turgidity	size, form and orientation of plant organs
cell structure	
senescence	distribution of fresh and dry biomass
phenology	phenology
leaf area index	row direction
soil pigmentation	soil roughness
soil moisture	soil moisture

λ=wavelenght; θ=incidence angle; ρ=polarization; ϕ=azimuth angle (look direction)

2.1.2.1 Radar operation

Airborne and spaceborne radar remote sensing uses an antenna mounted on an aircraft or spacecraft, actively transmitting a radar signal in very short bursts or pulses in a *side-looking direction* (SLR) towards the earth´s surface. By electronically measuring the return time of signal echoes, the range, or distance, between the transmitter and reflecting objects can be determined. As energy propagates in air at approximately the velocity of light (c), the slant range (SR) to an object is given by:

$$SR = \frac{ct}{2} \tag{2.1}$$

with SR=slant range, c=speed of light and t= time between pulse transmission and echo reception. The factor 2 is entered because the signal travels the distance to and from the target (Lillesand & Kiefer, 2000).

Imaging radar systems can be divided into two major categories: *Real Aperture Radar* (RAR) and the *Synthetic Aperture Radar* (SAR), depending on the imaging technique.

RAR
Real Aperture Radars transmit pulses from a side-looking antenna and are normally airborne, due to technical aspects concerning the antenna length. The spatial resolution in azimuth direction is dependent on the antenna footprint, being linearly proportional to the distance between sensor and surface target (Ulaby et al., 1982).

SAR
The synthetic aperture imaging technique in a SAR system uses the travelled distance of the sensor along track (azimuth direction) to simulate a larger antenna than it´s actual size. Whereby, the travelled length of the antenna between the pulses is the synthetic aperture. A single antenna moves along the track acquiring data, which has a similar effect as using an array of antennas.

Targets are illuminated several times from different locations along track, generating numerous echoes. These target echoes are than recorded coherently (amplitude and phase as a function of time) and finally combined to synthesize a linear array. By using a small antenna a higher spatial resolution can be achieved, independently of the sensor-target distance (Ulaby et al., 1982).

2.1.2.2 System parameters

The interactions between system parameters and target characteristics are essential for the interpretation of SAR data. SAR systems have specific operational characteristics, influencing the interactions between the transmitted pulses and the land surface targets. Frequency (wavelength), polarization and incidence angle are the primary system parameter for the definition of a SAR sensor and will be explained in the next section.

Frequency
Imaging SARs normally operate in a single band, which is defined by it´s frequency (wavelength). Reason for a single band operation is the limited power supply, due to the fact that SAR systems rely upon their own energy source. Whereby, high frequencies (short wavelengths) transmissions require a rather large amount of power, thus excluding their use in spaceborne systems. The microwave bands definitions are listed in Tab. 2.2, the used C-band is marked in bold (Ulaby et al., 1981).

The frequency of a sensor is important for the penetration and infiltration of the microwave signal into the material. The backscatter magnitude from agricultural targets is dependent upon frequency due to:

- differences in the dielectric constant of water as a function of frequency
- the relationship between frequency and plant part size and/or penetration depth

Whereby, shorter wavelengths and thus higher frequencies have a lower penetration. As a rule of thumb: the penetration is normally half of the wavelength.

Table 2.2: Microwave bands (Ulaby et al., 1981)

band	frequency	wavelength
P	< 390 MHz	> 76.9 cm
L	390 MHz - 1.55 GHz	76.9 - 19.35 cm
S	1.55 - 4.20 GHz	19.35 - 7.14 cm
C	**4.2 - 5.75 GHz**	**7.14 - 5.22 cm**
X	5.75 - 10.9 GHz	5.22 - 2.75 cm
K	10.9 - 36.0 GHz	2.75 cm - 8.3 mm
Q	36.0 - 46.0 GHz	8.3 - 6.5 mm
V	46.0 - 56.0 GHz	6.5 - 5.4 mm
W	56.0 - 100 GHz	5.4 - 3.0 mm

Agricultural targets are composed of significant and varying amounts of water, thus the frequency dependence of the dielectric constant is very important in the interaction process. With decreasing frequency (e.g., L- or P-band), the signal penetration into the crop/soil system increases and the sizes of the targets relative to the wavelength are smaller. In general, higher frequencies (e.g., X-band) are dominated by canopy scattering, while lower frequencies (e.g., L-band) have dominant soil backscatter contribution. Some studies have also pointed out that the combined use of different frequencies has a high potential for, e.g., biomass retrieval. Amodeo et al. (1996) conclude that L-band is also valid for biomass monitoring of crops with low density, while the use of L and C-band give useful information for high density crops. Another combined analysis was performed by Ferrazzoli et al. (2000). They studied different frequencies, polarizations and incidence angles during a whole crop growth cycle. They have figured out that at L- and C-band, the emissivity increases during the crop growth and decreases during crop drying. While at higher frequencies, it increases during crop growth and remains close to unity until the end of the cycle. They calculated frequency indices and analysed the correlation with wheat biomass during a crop life cycle. Their synopsis was that a two-frequency radiometer, operating at C- and X-band should be able to monitor wheat biomass during the whole growing phase.
These studies also indicate the potential of using new satellite sensors in a combined way, when acquisition time gaps are not to high, allowing to use the advantages of each frequency.

Polarization
Regardless of the frequency of the SAR system, SAR signals can be transmitted and/or received in different modes of polarization. The polarization refers to the direction of the electromagnetic wave. Microwaves are transverse, meaning that the vibrations are perpendicular to the direction of wave propagation. For SAR applications, waves are typically polarized in a plane, either horizontally (H) or vertically (V) polarized[1]. The polarization is defined for the outgoing (transmitted) and incoming (received) radiation. Thus, possible combinations are like-polarizations (VV or HH) and cross-polarizations (HV or VH). ERS-2 only possesses the VV mode (like polarization), Envisat ASAR also offers HH like polarization or modes like VV/HH, HH/HV or VV/VH. Microwave scattering from the land surface highly depends on the polarization of the wave transmitted. If the plane of polarization of the transmitted wave is parallel to the main line of the targets orientation the like-polarized backscatter is stronger (Henderson & Lewis, 1998).

Important in the agricultural context is the different attenuation of grain crops at HH versus VV. In general the vertical orientated components of crops (e.g., stems) couple much more effectively with vertically polarized signals, which results in a decreased backscatter (increased attenuation). Whereby, the HH polarization penetrates the crops more, thus reacting more with the soil moisture and soil roughness or the horizontally orientated leaves. This was also analysed by Mattia et al. (2003) using a ground-based scatterometer and is displayed in Fig. 2.4. The figure shows HH and VV polarizations acquired with different incidence angles for C-band over a crop growth cycle. For both analysed angles the VV polarization features a higher attenuation than the comparable HH polarization.
For many vegetation studies, the use of different polarizations, in particular cross polarization, will improve the discrimination between vegetation (volume scattering) and soil (surface scattering) and improves the retrieval of low biomass values[2].

[1] http://envisat.esa.int/handbooks/asar/CNTR.htm
[2] http://envisat.esa.int/handbooks/asar/CNTR.htm

The cross-polarization of the transmitted wave is depended on the amount of multiple scattering taking place at the monitored targets. Depolarization is defined as a change in polarization of the transmitted wave into the polarization orthogonal to this transmitted wave. It occurs as a result of significant multiple scattering within the target, as with very high surface roughness, or as a result of significant volume scattering from the vegetation canopy. Systems with cross-polarized receiving capabilities can also provide additional information and help understanding the target/wave interactions (Henderson & Lewis, 1998).

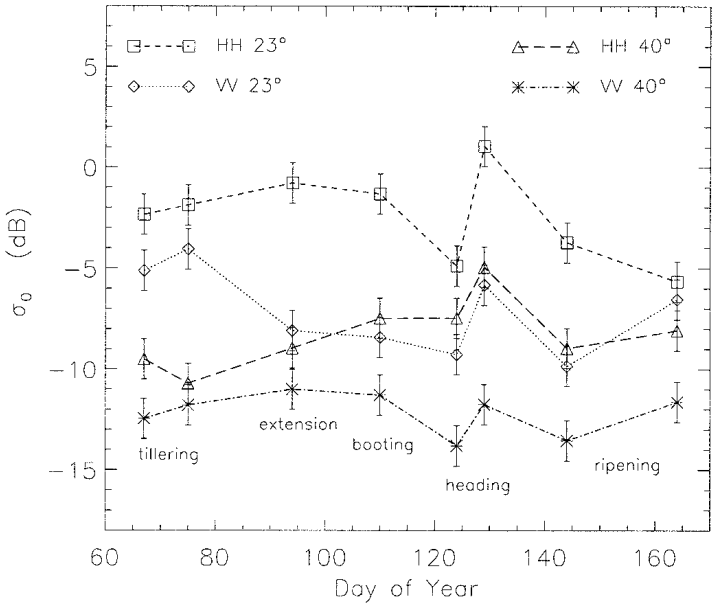

Figure 2.4: Multi-temporal backscatter values acquired at HH and VV polarizations and at 23° and 40° incidence angle (C-band) (Mattia et al., 2003).

The multi-temporal pattern will be explained in a later section.

Incident angle

The incidence angle (θ) is defined as the angle formed by the SAR beam and a line perpendicular to the surface at the point of incidence[3]. Fig. 2.5 schematically illustrates the system and the local incident angles. Whereby, in a flat surface the incident angle (θ) is complementary to the depression angle (γ) of the sensor system. Generally, smaller incidence angles (θ) result in higher backscatter, only for very rough surfaces the backscatter is not dependent of the incidence angle. The surface roughness, which will be discussed later in the section target characteristics, changes in dependence of the local incidence angle.

Fig. 2.6 shows the angular variation (23°, 30°, 40°, 50° and 60°) of wheat backscatter (VV) measured with a ground-based scatterometer (Mattia et al., 2003). They point out, that during early phenological stages (tillering) the VV backscatter at 23° is very close to the backscatter from bare soils, while at higher incidence angles the VV backscatter is significantly lower than the measured bare soil backscatter. This means that at high incidence angles the soil backscatter component is significantly reduced even at early growth stages, were the plants are just 0.10 -

[3] http://envisat.esa.int/handbooks/asar/CNTR.htm

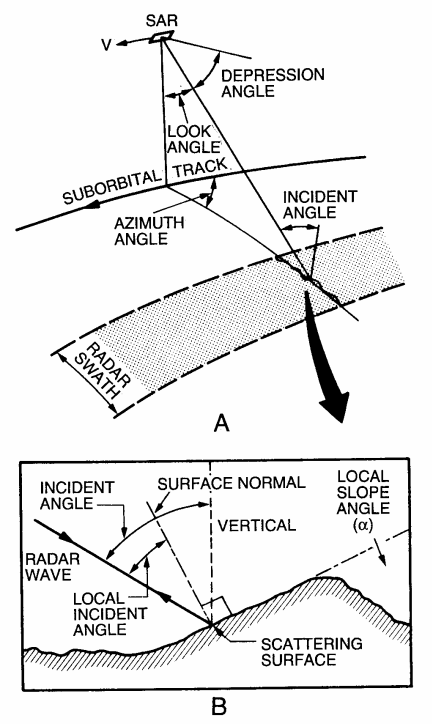

Figure 2.5: Radar geometry
Schematic graphics of the system (A) and local (B) incident angles (Henderson & Lewis, 1998).

0.15m high. With wheat growth (0.20 - 0.30m) the backscatter at 23° also decreases due to the soil attenuation. But here also an increasing attenuation from low to high incidence angles can be observed. From heading to ripening there is a moderate backscatter decrease from 23° to 40°, indicating that there is already a significant attenuation at low incidence angles. The increase after 40° indicates a change in the dominant scattering mechanism from soil backscatter to canopy backscatter.
In summary these figures also underline the general assumption that higher incidence angles are more suitable for vegetation monitoring, due to the lower soil contribution. But it also shows that during a wheat growing season the angular response changes before and after heading. In the 23° to 40° range, the angular dependence suggests that the dominant scattering mechanism is the soil scattering, attenuated by the canopy. While at higher incidence angles the dominant mechanism change to canopy scattering after heading (when a change in the plant geometry occurs).

2.1.2.3 Target characteristics

Interactions of the microwaves with agricultural targets are significantly influenced by the system parameters, such as geometry (and the associated roughness) and with the target characteristics such as moisture content (and associated dielectric constant) of crops and soils as well as their inter-relationships. In the first section the backscatter coefficient will be introduced, as it is a quantitative measure of backscatter intensity. Surface roughness and electrical characteristics of the land surface targets will be described afterwards.

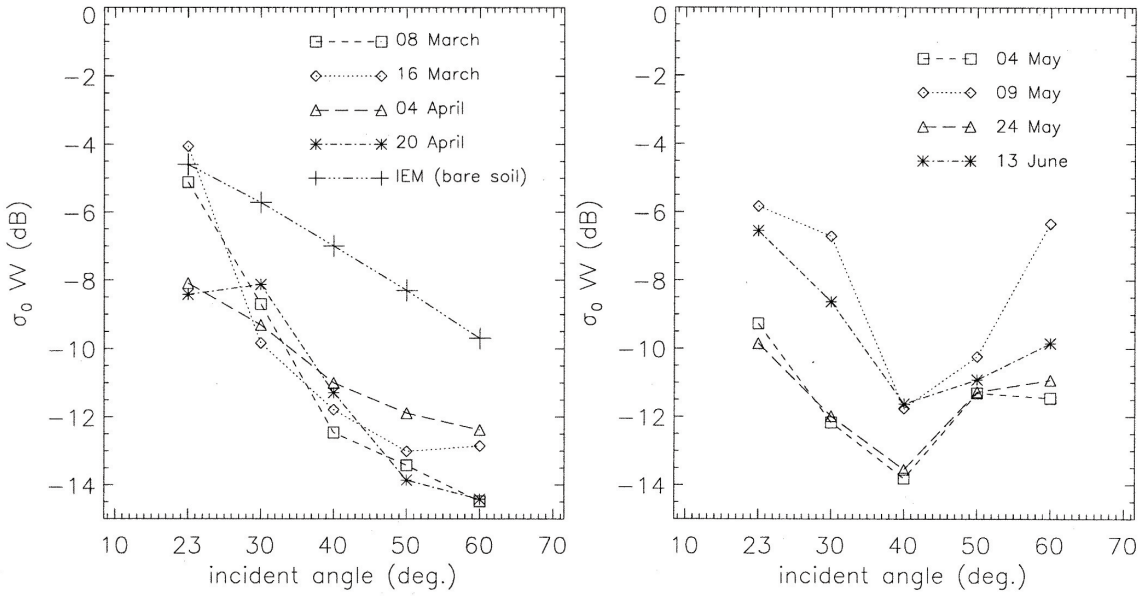

Figure 2.6: Backscatter values (VV) versus incidence angles during a phenological wheat cycle (Mattia et al., 2003).

Backscatter coefficient

Targets generally scatter the energy transmitted by the SAR in all directions. Backscatter is the portion of the transmitted SAR signal that is redirected backwards directly by the target towards the SAR antenna. Whereby, the intensity of each pixel is proportional to the ratio between the density of energy scattered and the density of energy transmitted from the targets on the land surface (Waring et al., 1995).

The backscattered energy depends on the variable SAR cross-section (σ), which is the amount of transmitted power absorbed and reflected by the target. It is a measure of the reflective strength of a SAR target. The normalized measure of the SAR return is called the backscatter coefficient or sigma naught (σ^0), defined per unit on the ground. Sigma naught is a characteristic of the scattering behaviour of all targets within a pixel. It varies over several orders of magnitude and is normally expressed as a logarithm with decibel units.

The return power is given by (Moore, 1970):

$$P_{pq} = \int\int_{A_0} \frac{P_t G_r G_t \lambda_2}{(4\pi)^3 R^4} \sigma^0_{pq} ds \tag{2.2}$$

with P_{pq}=transmitted power polarization dependent; p=received; q=transmitted polarization
A_0=illuminated area
P_t=transmitted power
G_r=receiver-antenna power gain
G_t=transmitter-antenna power gain
λ=incident wavelength
R=range between target and radar
σ_{pq}=radar scattering cross-section of the target; polarization dependent

It is the generalized return power equation of a radar, also taking into account differential targets, the illuminated area and the polarization dependence.

Generally σ^0 is a function of system parameters (i.e. frequency, polarization and look angle) and target parameters (i.e. surface roughness and permittivity) (Ulaby et al., 1982).

The speckle or salt-and-pepper-effect in the SAR images is an interference phenomenon produced between backscatter coming from many targets within a pixel. Speckle represents true electromagnetic scattering and influences the interpretation of SAR images. In chapter *DATA* (3) a method for speckle reduction will be shown. Speckle filtering is an important preprocessing step, when working with SAR data.

Surface roughness
Surfaces roughness is an important target characteristic influencing the strength of the backscatter and has to be considered in relation to the scale of the targets observation. Three scales are often divided: microscale, mesoscale and macroscale roughness (Lillesand & Kiefer, 2000; Brisco & Brown, 1998).

Microscale roughness refers to rather small targets within an individual pixel, i.e. leaves and branches of trees or stones. Microscale roughness is measured in cm and is dependent on wavelength, depression angle target height. Normally the modified Rayleigh criterion is used for expressing this relationship (Jensen, 2000):

$$h < \frac{\lambda}{25sin\gamma} \tag{2.3}$$

with h the local height of target, λ the wavelength (cm) and γ the depression angle (°). For example, computing h using $\lambda = 3cm$ (X-band) and $\gamma = 45°$, result in $h < 0.17cm$. Meaning, that if the local height of a target is below 0.17 cm the target´s surface is considered smooth and as a near-perfect specular reflector.
Other microscale scattering mechanisms are shown in Fig. 2.7. These are the *double bounce* or corner reflector scattering occurring often in urban areas, the *double bounce* from two natural smooth surfaces like grass and freshly cut tree stumps. Complex scattering forms are *volumetric scattering* in trees or other plants or the *volumetric scattering* in ice and snow. Macroscale surface roughness is influenced by topographic slope and the aspect of the terrain, causing shadows. These patterns are normally many times larger than an individual pixel.

Dielectric constant
Electrical characteristics of terrain features determine next to the geometric features the intensity of the backscatter. The *complex dielectric constant* is a parameter indicating the reflectivity and conductivity of various materials. Whereby, with increasing reflectivity and conductivity the value of the complex dielectric constant increases (Lillesand & Kiefer, 2000).
Most natural materials have a dielectric constant between 3 to 8 when dry, whereas water has a dielectric constant of around 80 in the microwave region of the spectrum. Meaning that the presence of water (moisture) in soils or vegetation will significantly increase the reflectivity.
For soils this implies that the presence of water in the top few centimetres of bare soil can be detected in SAR imagery, becoming particularly apparent at longer wavelengths.
The vegetation canopy interacts with SAR waves as a group of volume scatterers. The canopy is also composed of a large number of plant components (leaves, stems, stalks and so on) with different dielectric constants. Additionally, the canopy has a subjacent soil, resulting in surface scattering of energy that penetrates the vegetation canopy.

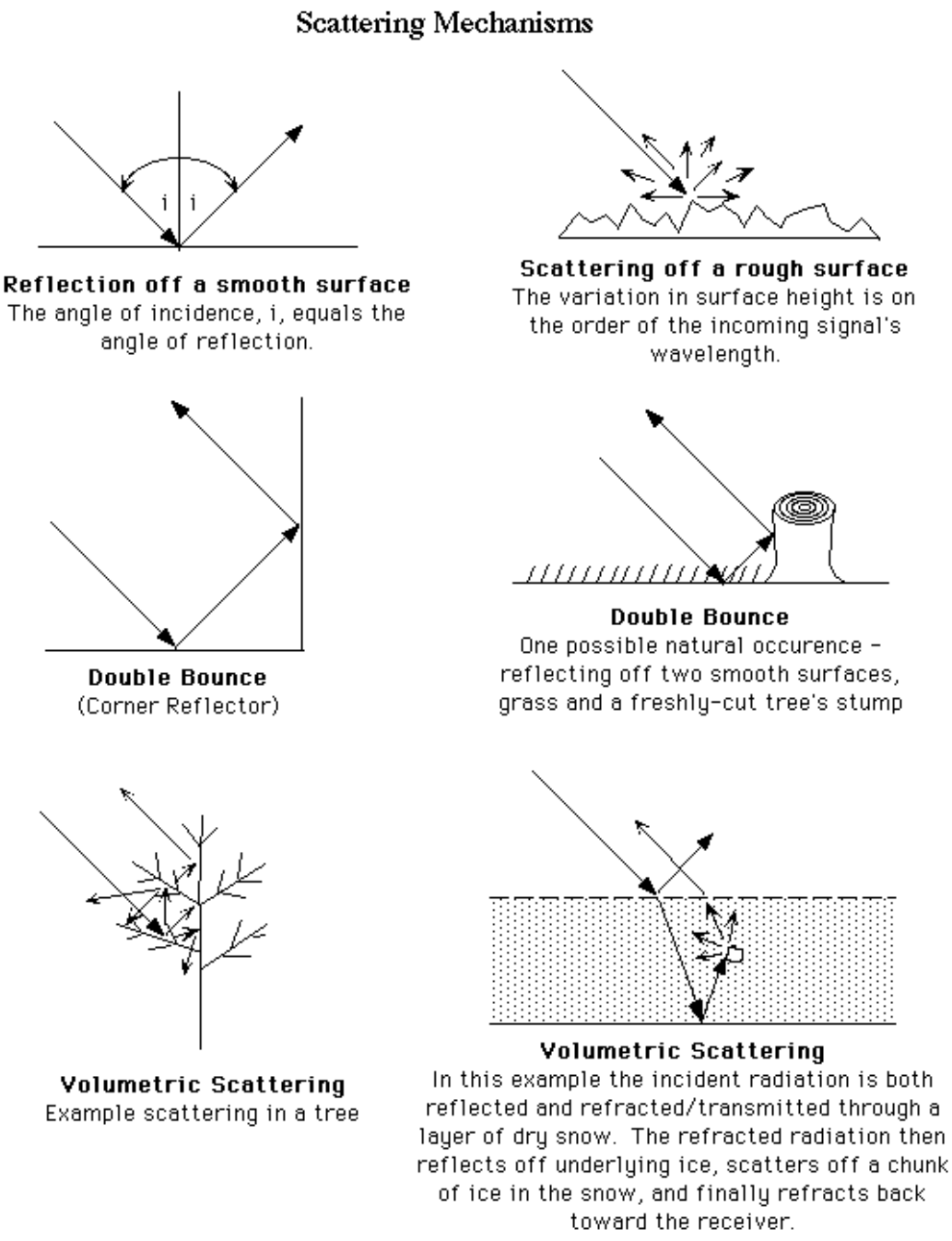

Figure 2.7: Different scattering mechanisms

http://envisat.esa.int/handbooks/asar/CNTR.htm

2.1.2.4 SAR image characteristics

SAR images can be configured either in slant range or ground range. Slant range represents the distance measured along a line between the radar antenna and the target. Image direction as measured along the sequence of line-of-sight rays from the radar to all reflecting points in the illuminated scene[4]. The spacing between return signals on slant range imagery is directly proportional to the time interval between echoes from adjacent terrain features. Ground range is the perpendicular distance from the ground track to a given object of the land surface. Also defined as the range direction of a side-looking radar image as projected onto the nominally horizontal reference plane. Ground range projection requires a geometric transformation from slant range, which can lead to geometric distortions on the radar images, due to the side-looking geometry of the systems (Ulaby et al., 1982).
The mentioned geometric distortions are (see Fig. 2.8) *foreshortening*, where the fore slopes appear to be compressed; *layover* or the reverse ordering of surface elements, resulting in imaging the tops of objects before their bases (most serve at near range with steeper incident angles); *shadow* caused by slopes facing away from the radar antenna with an angle that is steeper than the radar depression angle, thus returning weak signals or no signals at all (Lillesand & Kiefer, 2000).

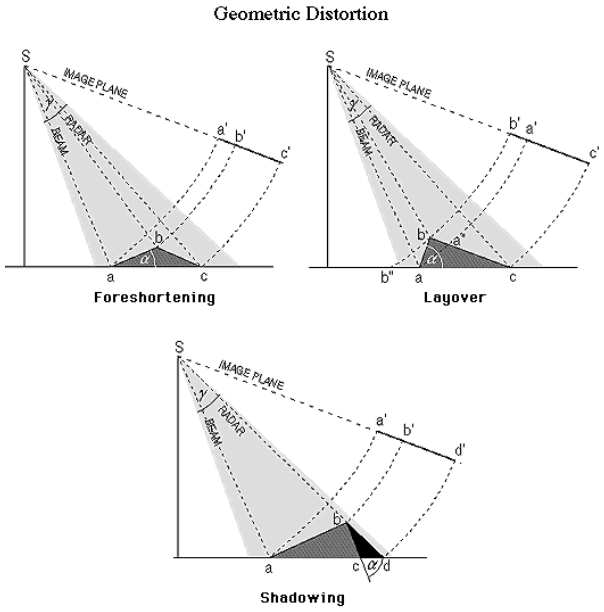

Figure 2.8: Effects of geometric distortions in SAR images (Lillesand & Kiefer, 2000).

2.1.3 Multi-temporal backscatter characteristics during the crop growth cycle

Vegetation canopies can be divided into several groups, depending on the complexity of the canopy architecture and the size of the scattering elements, relative to the wavelength (Ulaby et al., 1986). The next section will focus mainly on crops with small leaves like winter wheat.

[4]http://envisat.esa.int/handbooks/asar/CNTR.htm

Fig. 2.9 shows all the primary factors controlling SAR backscatter from the land surface and their interactions. It is a complex subject and a challenge for applications extracting information about vegetation biomass. As diurnal variations in σ^0 can have various reasons and the relationship changes during a phenological cycle. It is also important to note that there is also a dependence on the system characteristics, e.g., wavelength, incidence angle, polarization.
In a generalized form the backscatter modelling of a canopy is done by considering:

$$\sigma^0 = \sigma_{vegetation} + \tau^2 \sigma_{soil} \tag{2.4}$$

Thus, considering backscattering interactions from the vegetation and the soil, which is influenced twice. Problematic hereby is to adequately describe the canopy, due to it´s complex and variable geometry. And to define the soil term, due to it´s changing impact during crop growth and the changing two-way transmissivity through the vegetation layer. These terms and all changes during crop growth, cause a typical temporal backscatter pattern.
Fig. 2.10 shows the typical multi-temporal backscatter pattern of wheat. The generalized backscatter pattern of cereals has rather high values during spring, with increasing plant growth the backscatter decreases again, until May/July when it experiences a reflection minimum and thus a absorption maximum. Afterwards, the backscatter increases again with the crop ripening. However, it is important to keep in mind that the weather situation (wind or rainfall) also influences the backscatter signal.
Wooding et al. (1995) and Bouman & van Kasteren (1990) assume that due to wind influences the plant geometry changes and thus the backscatter is influenced. Whereby, especially for small plants with thin stems as wheat, with vertical leaves dominating the backscatter, the impact is rather high. Also the ear orientation seems to be effected by wind, especially when the stems of the ears are bent and the ears lay nearly horizontally in the top of the canopy. Thus, it is hard to quantify the real influence of wind direction and wind strength.
Rainfall actually has a direct effect due to moisture changes within the canopy and on the leaves and an indirect effect caused by increasing soil moisture (Bouman & van Kasteren, 1990). Dew has an impact, due to an increased dielectric constant, whereby the influence is dependent on crop type, frequency and polarization (Riedel & Schmullius, 2003; Wood et al., 2002). Thus, it is important when comparing data from ascending and descending orbits, whereby the descending orbit (acquisition time in the morning) can be effected by dew, having higher backscattering coefficients. If images from both orbits are combined, it may be difficult to separate dew effects from target changes.

2.1.3.1 Early growth stages: germination-tillering

During the early growth cycle of cereals the backscatter signal is dominated by bare soils, producing mainly surface scattering. The governing factors are soil surface roughness and soil moisture content (Ulaby et al., 1986). Thus, the small vegetation cover plays a minor role. Fig. 2.11 shows the SAR sensitivity to soil moisture as a function of frequency at various angles of incidence. With increasing frequency (shorter wavelength) the sensitivity of the sensor regarding soil moisture decreases. The C-band (4.2 - 5.75GHz) used is rather sensitive. Also illustrated is the impact of varying incidence angles, whereby the sensitivity decreases with increasing angle, e.g., ERS-2 having an incidence angle of around 23.5°.

Fig. 2.12 displays the measured scattering coefficient as a function of soil moisture for three different surface roughnesses. With increasing soil moisture the backscattering coefficient in-

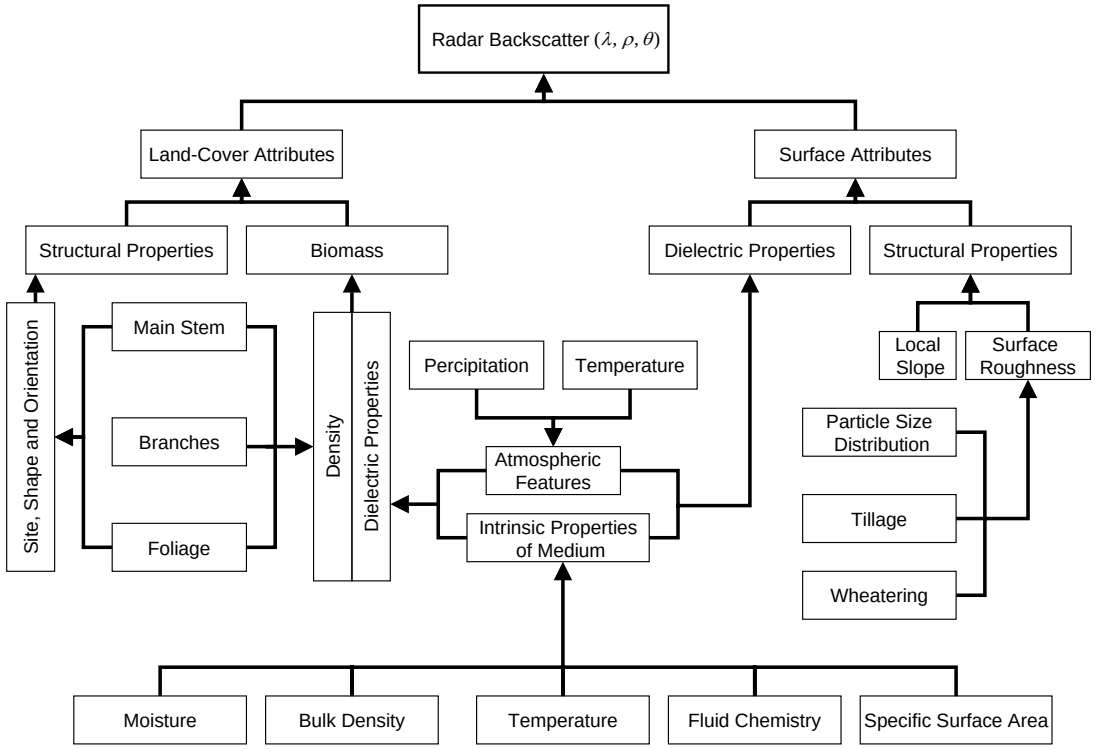

Figure 2.9: Primary factors controlling the SAR backscatter, modified after: Dobson et al. (1995).

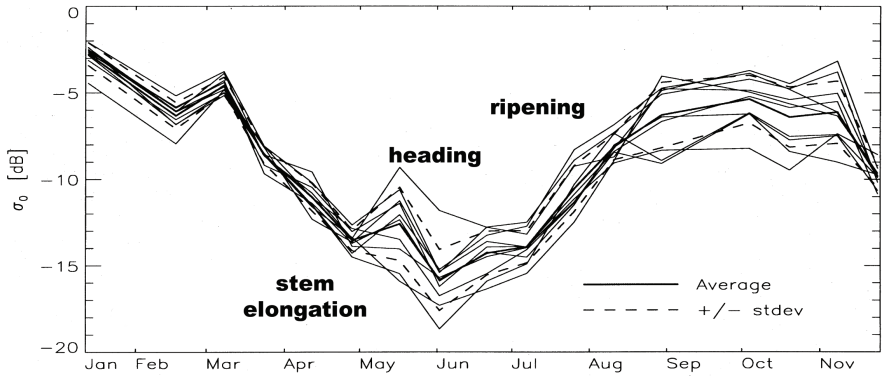

Figure 2.10: Temporal backscatter pattern of ten winter wheat fields, modified after: Borgeaud et al. (1995).

creases, whereas the surface roughness plays a more marginal role. In natural environments the surface roughness also decreases due to weathering.

2.1.3.2 Middle growth stages: stem elongation-heading

With progressing crop development, volume scattering within the plant canopy increases, becoming the dominant scattering mechanism. This leads to a continuous increase of backscatter values in the cross-polarization, due to depolarization of the microwaves at the coincidentally orientated vegetation components (Ulaby et al., 1986). During stem elongation the canopy

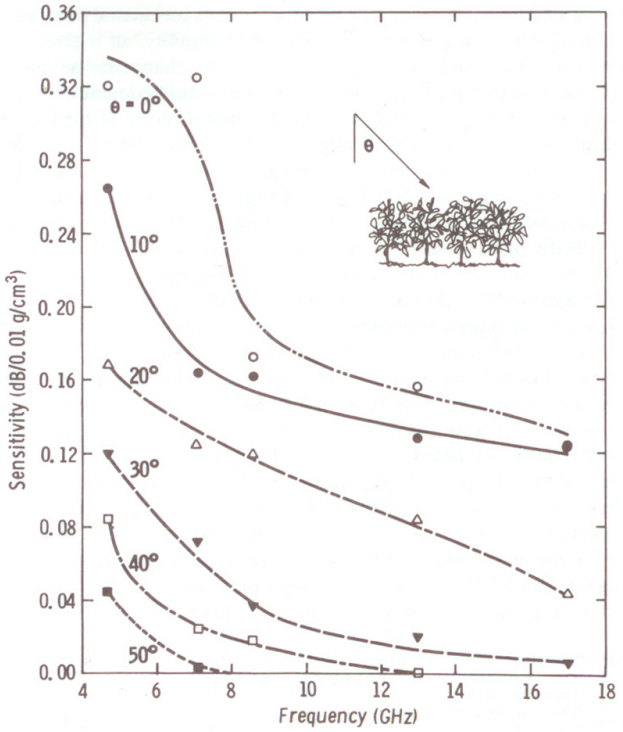

Figure 2.11: SAR sensitivity to soil moisture in dependence of frequency and incidence angle (Ulaby et al., 1986).

changes from an horizontal structure to a stalk-dominated vertical structure. The vertical oriented components of grain plants couple much more effectively with vertical polarized waves, resulting in increased absorption and lower backscatter. Another feature reducing the backscatter is mentioned by Schmullius et al. (1993) and Schmullius & Nithack (1995) is that due to tillering a change of vegetation surface roughness happens. From the broad primary leaves (high backscatter) to elongation of the plant. Another feature to mention is that crop phenology governs plant water content and thus the crops´s dielectric properties (Brisco & Brown, 1998). As crop growths the water content increases, which generally increases the contribution to σ^0 from the plants and decreases the penetration of the transmitted microwave energy into the vegetation. This then decreases the soil contribution to the total backscatter signal (Ulaby & Wilson, 1985).

2.1.3.3 Late growth stages: anthesis-senescence

During ripening the backscatter signal increases again caused by multiple effects. Due to decreasing plant coverage (less water content) the soil contribution to the backscatter increases again. Spoenemann & Schieche (1997) explain the increase with a changing plant geometry. The canopy becomes sparser, reducing the volume scattering and the soil contribution increases. Ear appearance is another factor, having a strong contribution to the backscatter (Siqueira & Sarabandi, 1996). But all these single effects are rather hard to quantify.

Hamacher (2000) has used this multi-temporal backscatter pattern (VV polarization) for extracting the important heading/flowering date of winter wheat (see Fig. 2.13). What can be observed

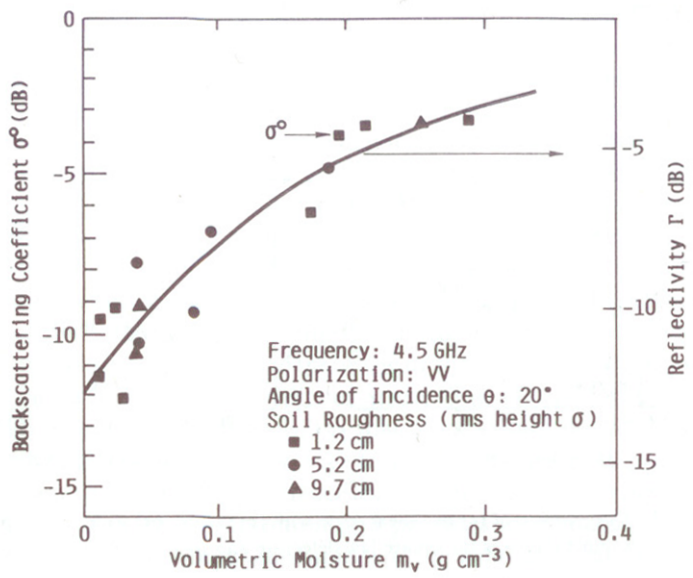

Figure 2.12: Backscattering coefficient as a function of soil moisture content for three surface roughnesses (Ulaby et al., 1986).

in the idealized curve is a rather high backscattering coefficient in spring, which decreases with plant growth until heading/flowering, afterwards an increase until senescence occurs, due to the effects just explained. The heading/flowering date is marked trough a reflection minimum -absorption maximum- and can be easily extracted by multi-temporal analysis of VV polarized data.
The heading/flowering date is an important stage for crop growth modelling, because after reproductive growth begins, the number of kernels per plant is often calculated from plant weight at flowering and kernel weight is calculated from time and temperature available for dry matter distribution, respectively. Final yield is estimated as a product of plant number, kernel number, and kernel size (Rickman & Klepper, 1991). Thus, it is an important calibration and validation parameter for crop growth modelling.

2.2 Retrieval of vegetation features from earth observation data

Remote sensing is a very powerful method for providing information about the vegetation in various wavelength. In order to retrieve the desired information or to relate the canopy state variable to the remote sensing information there are different kinds of models: deterministic, semi-empirical or empirical (Atzberger, 2003; Guissard et al., 2005; Moulin et al., 1998).
The complex deterministic models are based on the inversion of radiative transfer models, considering plant canopy and the underlying soil. These models are efficient techniques for describing the expected signal in the different frequencies based on variable parameters, e.g., soil properties, leaf structure or observation geometry. Well known models for multispectral and hyperspectral data are the coupled SAIL (Scattering Arbitrary Inclined Leaves) (Verhoef, 1984) and PROSPECT (model of leaf optical PROperties SPECTra) (Jacquemoud & Baret, 1990)

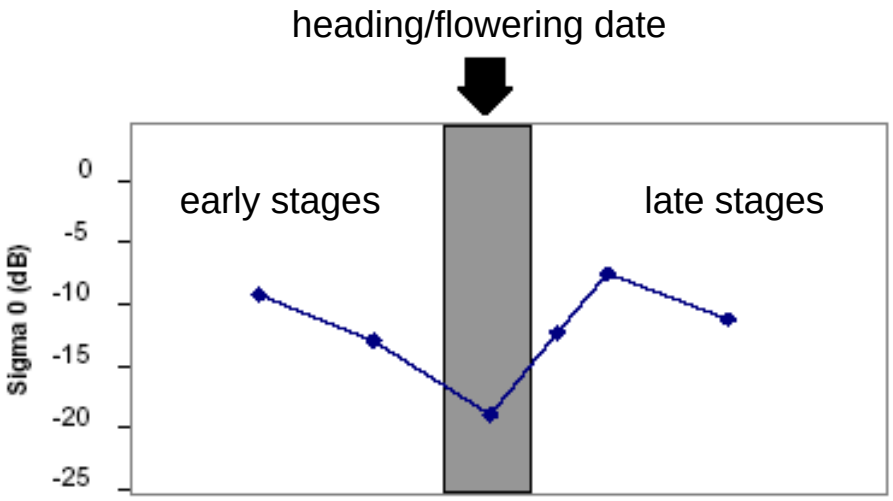

Figure 2.13: Idealized backscatter curve (winter wheat) with the heading/flowering date marked, modified after: Hamacher (2000).

models. Whereby, the SAIL model calculates the directional reflectance on the upper layer of the canopy, using a function of structural and spectral properties of the vegetation canopy and PROSPECT provides the reflectance and transmittance of fresh leaves over the whole solar domain. Frequently used radiative transfer models for modelling the backscatter are the second-order radiative transfer model (RT2) (Saich et al., 1995) and the first-order Michigan Microwave Canopy Scattering Model (MIMICS) (Ulaby et al., 1990). Difference between both models is that RT2 also predicts double scattering contributions from the canopy. Problematic with these complex models is that they not allow an analytical inversion, thus other inversion techniques like look-up-tables (LUT), numerical inversion, neural networks or predictive equations have to be used (Atzberger, 2003). Also problematic with the deterministic models is the large amount of necessary input variables, making it hard to transfer the findings to regional analysis.

The semi-empirical models combine a physical model, whereby for the model inversion empirical relationships are used. Probably the best known semi-empirical model is the approach by Monteith (Monteith, 1977) for estimating biomass using remote sensing data. The assumption hereby is, that healthy plants with sufficient water and nutrition supply have a netto photosynthesis and thus biomass production proportional to the absorbed photosynthetic active radiation (PAR). Other models are, e.g., the CLAIR model for multispectral data (Clevers, 1989) (section 2.2.2) and the Water Cloud Model for SAR data (Attema & Ulaby, 1978) (section 2.2.3) used within the presented thesis.

Empirical models are based only on the statistical analysis of experimental data. Using ground truth data, transfer functions between the desired biophysical parameter and the remote sensing data are modelled. This method is rather simple, but has the disadvantage that the calibrated transfer functions are hard to transfer to other regions or within time. Classically used relationships are, e.g., the linear relationship between the NDVI and LAI (Wiegand et al., 1992) or the exponential and linear functions between backscatter and LAI (Brakke et al., 1981; Mattia et al., 2003), which are also analysed within this study.

2.2.1 Statistical regression considerations

For a better understanding of the derivation of biophysical parameters some statistical considerations have to be introduced, for a detailed statistical introduction please refer to, e.g., Schoenwiese (1992).
Major purpose of regressions is to study the dependence of one set of variables to another. Thus, in the remote sensing context to analyse the relationship between the received signal and, e.g., a biophysical parameter. In a simple linear regression the relationship is assumed to be of the form:

$$E(Y/X = x) = \alpha + \beta x + \epsilon \tag{2.5}$$

with Y the random variable to predict, X the random variable Y is predicted from, α and β coefficients, which are fixed but unknown and ϵ the error term. The specification is that Y is the dependent variable, e.g., the biophysical parameter and X the independent variable, e.g., the vegetation index. It can be argued that the remote sensing signal received, is dependent on vegetation status. This, in the remote sensing literature known "specification problem" is taken into account. The other problem is that it assumes no error in measurement of vegetation reflectance.
In multiple regressions the relationship between several independent (predictor) variables and a dependent variable are analysed. E.g., the SAR backscattering coefficient is not only dependent on the vegetation parameter, but also influenced by soil characteristics.

$$E(Y/X_n = x_n) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 ... + \beta_n x_n + \epsilon \tag{2.6}$$

The more general non-linear regression equation is:

$$E(Y/X = x) = k(x; \theta) \tag{2.7}$$

where k is an arbitrary function, X is a random vector, x is a corresponding vector of fixed values and θ is a vector of parameters.
For determining the values of the unknown coefficients the sum of the residuals (difference between the predicted and observed values) are minimized. This method is knows as ordinary least squares.
For assessing the strength of a relationship between random variables there are different correlation coefficients. Probably the best known is the *Pearson product-moment correlation coefficient* R, defined as the sum of the products of the standard scores of the two measures divided by the degrees of freedom:

$$r = \frac{\Sigma z_x z_y}{n - 1} \tag{2.8}$$

The coefficient ranges from -1 to 1, whereby values of 1 show a perfectly positive description of the relationship. This coefficient is restricted to 2-dimensions linear relationships.
The coefficient of determination R^2 is another measure for assessing the proportion of variability in a data set. Whereby, variability is defined as the sum of squares.

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \tag{2.9}$$

with SS_T as the total sum of squares, SS_R the regression sum of squares and SS_E the sum of squared errors. The coefficient ranges from 0 to 1, with 1 indicating that the regression line perfectly fits the data. The advantage of using R^2 is that it is also applicable for non-linear relationships and for more than 2-dimensions.

After calibrating a regression, thus analysing the relationship and defining the unknown coefficients, a validation is necessary. This validation has to be done on the basis of an independent data set. Thus, normally one divides the total data set into a data stack for calibration and one for the independent validation of the constructed model. If this is not possible, due to the amount of samples the cross-validation or rotation estimation can be used. This is a practice of partitioning a sample of data into subsets such that the analysis is initially performed on a single subset, while the other subsets are retained for subsequent use in the validation process. The used leave-one-out cross-validation method uses a single observation from the original sample as the validation data and the remaining observations as the training data. This is repeated until each sample is once used as the validation data.
The error estimation is computed as the mean average error and the root mean square error (rmse). The mean average error is actually a weighted average of the absolute error from a single cross-validation run.

The root mean square error (rmse) is a measure of the difference between values predicted by a model and the actually observed values, the residuals. It is defined as the square root of the mean square error (mse):

$$rmse(\hat{\theta}) = \sqrt{mse(\hat{\theta})} = \sqrt{E(((\hat{\theta}) - \theta)^2)} \tag{2.10}$$

with $\hat{\theta}$ as an estimator.

2.2.2 Biophysical parameter retrieval using multispectral data

Within the following section different possibilities of retrieving biophysical parameters from multispectral remote sensing data are described. Starting with a rather easy approach using empirical modelling to the more sophisticated semi-empirical model CLAIR. Biophysical parameters of interest are hereby the fresh and dry biomass of winter wheat, as well as the plant water content and the leaf area index. These are all important parameters characterising the plants vitality and the phenological stages.

2.2.2.1 Empirical modelling

Empirical models are based on the statistical analysis of the experiment data, e.g., ground truth data and the remote sensing data. Whereby, the multispectral vegetation profile is often transferred into vegetation indices (VI), highlighting important vegetation characteristics and trying to minimize the soil and atmosphere influence. They normally reduce the multi-band observation to a single numerical index, typical being a sum, difference, ratio or other linear combination from two or more wavelength intervals. Generally, there is the question, which VI should be used, because each features advantages and disadvantages. In Baret & Guyot (1991) a couple of VIs and their performances were compared, which were NDVI (*Normalized Difference Vegetation Index*) (Rouse et al., 1978), PVI (*Perpendicular Vegetation Index*) (Jackson, 1983) and the SAVI (*Soil Adjusted Vegetation Index*) (Huete, 1988). The last two VIs consider the soil signal and are appropriate for early growth stages with low vegetation coverage. For the middle

growth stages (middle vegetation coverage) all VIs exhibit a similar behaviour. For dense vegetation coverage the NDVI performs best.
The NDVI is a simple numerical indicator and most commonly used in the remote sensing community for assessing healthy green vegetation, using the near infrared and visible wavelengths (Rouse et al., 1978). Healthy green plants absorb solar radiation in the photosynthetically active radiation (PAR) spectral region, which is their source of energy for the photosynthesis process. On the other hand leaf cells scatter (e.g., reflect and transmit) solar radiation in the near-infrared spectral region. The energy level per photon in that domain would result in overheating the plant and possibly damage the tissues when absorbed. Hence, healthy green plants exhibit rather high NDVI values, while diseased vegetation or non-vegetated areas feature rather low or even negative NDVI values (e.g., water).
Various studies have already assessed the relationship of VIs to different biophysical parameters (e.g., crop biomass, chlorophyll content or leaf area index) (Ustin, 2004). Wiegand et al. (1990) gave an overview over some regression analysis and measured coefficients of determination, rms errors and residuals for retrieving LAI using different vegetation indices. The R^2 within their study all vary between 0.71 using a linear NDVI regression and 0.93, when using the greenness vegetation index and a quadratic relationship. Thus, in general in all studies very strong relationships were found.
But these studies also point out the disadvantage of using VIs, because they saturate at high LAI values. Also causing problems when interpreting the VIs is the soil signal, which especially in arid regions or during early phenological stages influences the remote sensing signal.

2.2.2.2 Semi-empirical modelling

Due to all these negative effects of the simple empirical modelling using a VI the simplified, semi-empirical reflectance model CLAIR was designed (Clevers, 1988, 1989; Clevers et al., 2002). In this model, first of all the *W*eighted *D*ifference *V*egetation *I*ndex (WDVI) is ascertained. It is defined as a weighted difference between measured near-infrared and red reflectance, whereby a correction of the soil background influence is performed:

$$WDVI = NIR - (C \times RED) \tag{2.11}$$

With RED and NIR notating the spectral reflectance measured in the respective spectrum and C as the slope of the soil line, or the ratio between NIR and RED reflectance of soil. Especially for multi-temporal analysis a background (soil) correction has to be made, when ascertaining the relation between reflectance and a biophysical parameter. During the phenological cycle first the juvenescent plants have a low soil cover, later the green plants totally cover the soil until the end of the growing season, when plants show signs of senescence and turn to yellow.
In a second step the relationship between the biophysical parameter and the WDVI is compounded by:

$$biophysical\,parameter = -\frac{1}{\alpha} ln \left(1 - \frac{WDVI}{WDVI_\infty}\right) \tag{2.12}$$

The parameters α and $WDVI_\infty$ have to be estimated empirically from training sets, but they have a physical nature. It is a special case of the inverse of the *Mischerlich* function. In the special situation the curve runs through the origin and thus only has two parameters (Clevers, 1988). In Clevers (1988) the model was proposed for the derivation of the LAI, but because actually the LAI is also correlated with the parameters dry biomass, fresh biomass and plant

water content it should be possible to use the model for these parameters as well.

2.2.3 Biophysical parameter estimation using SAR data

As for the multispectral data within this section first biophysical parameter retrieval using rather simple empirical modelling approaches and afterwards using the semi-empirical Water Cloud Model (WCM) will be described.

2.2.3.1 Empirical modelling

Several studies have already analysed the relationship between the backscatter and vegetation parameters by using regression analysis (Amodeo et al., 1996; Brakke et al., 1981; Bouman & van Kasteren, 1990). Non-linear regression analysis were used by Bouman & van Kasteren (1990) analysing X-band SAR. They tried to estimate crop parameters by second order polynomial or logarithmic equations and achieved coefficients of correlation (R) for wheat between -0.48 (crop water content) and -0.84 (crop height). Amodeo et al. (1996) tried to monitor vegetation features using multi-temporal and multi-frequency (P-, L- and C-band) data and a regression analysis by adopting linear relationships. Their results showed that for wheat there is almost no relationship at P-band (R^2=0.18), a moderate one was noted at L-band (R^2=0.66) with a further improvement at C-band (R^2=0.75). Mattia et al. (2004) analysed the relationship between HH/VV ratio and LAI by using a simple power law and achieved very good correlations (R^2=0.81). But they also mention some problems, namely that often only a relative small number of rather homogeneous fields is analysed. Thus, a generalization of the analysed relationship over fields with different soil surfaces or planting densities is not possible. Another point is that the theoretical modelling is still lacking, thus the role of soil condition (soil moisture and soil roughness) especially for variable incidence angles is not clear.

2.2.3.2 Semi-empirical modelling

Attema & Ulaby (1978) developed the *W*ater *C*loud *M*odel (WCM) assuming that the microwave dielectric constant of dry vegetation is smaller than the dielectric constant of water, and because a vegetation canopy is composed of more than 99% air by volume, that the canopy can be modelled as a water cloud, whose droplets are held in place by the vegetation matter. The droplets within the canopy cloud are randomly distributed. By integrating scattering and attenuation cross-section contributions of N droplets per unit volume over the signal pathlength through the canopy, an expression was derived for the backscattering coefficient as a function of: volumetric soil moisture, volumetric vegetation moisture and plant height:

$$\sigma^0 = C[1 - exp(-Dwh/cos\theta)]cos\theta + A\, exp(Bm_s - Dwh/cos\theta)cos\theta \tag{2.13}$$

With σ^0 denoting the backscattering coefficient in dB, A,B,C,D the model coefficients, w the water content per unit volume [kg/m^3], θ the incidence angle [°radian], m_s the volumetric soil

moisture content [kg/m^3] and h the plant height [m]. The four coefficients needed correspond to different features: A and B to vegetation, whereby A corresponds to the albedo of the vegetation and B is an attenuation factor. Coefficients C and D correspond to soil features, whereby C can be considered as a calibration constant and D is the sensitivity of the signal to soil moisture.

For keeping the model rather simple, the following assumptions were made:

1. The cloud representing the vegetation consists of identical water particles, uniformly distributed throughout the space.

2. Only single scattering needs to be considered.

3. The only significant variables are cloud height and cloud density, which are assumed proportional to the volumetric water content.

In a generalized form the model can be written for the whole canopy as:

$$\sigma^0 = \sigma^0_{veg} + \tau^2 \sigma^0_{soil} \tag{2.14}$$

The SAR cross-section of the canopy σ^0 for an incidence angle is hereby expressed as the incoherent sum of the contribution of the vegetation layer (σ^0_{veg}), the contribution of the soil (σ^0_{soil}) and the two-way attenuation through the vegetation (τ^2). In their study Attema & Ulaby (1978) analysed the relationship over a period of several months, incidence angles (0-70°) and frequencies (8-18GHz) for HH and VV polarization and achieved correlation coefficients (R) ranging from 0.7 to 0.99.
Prevot et al. (1993) introduced a simple parametrization of the angular effect of soil roughness allowing to process multi-incidence angle data, which is especially of interest when working with Envisat ASAR or even when analysing different ERS-2 tracks. Another positive aspect is, that when the density of vegetation increases, volume scattering cannot be neglected any more and it‘s angular effects are similar to those induced by an increasing soil roughness, which might be confusing. Main aim of their approach was to keep the model simple, because one of the advantages of the WCM is it´s easy handling and inversion for biophysical parameter retrieval. Therefore they introduced the angular effect as followed: several theoretical or experimental studies have shown that for frequencies greater than 4GHz, the sensitivity of D ($\sigma^0_{soil} = C + Dm_s$) ($ms$= soil moisture) can be assumed to be constant in a first approximation. The more important dependence with the SAR configuration and the soil roughness leaves therefore the parameter C. The variations of C with the incidence angle can be linearised representing an averaged behaviour of the SAR cross-section in terms of roughness effects by:

$$\sigma^0_{soil} = C(F, \theta) + Dm_s \tag{2.15}$$

with

$$C(F, \theta) = C_1(F) - \theta C_2(F) \tag{2.16}$$

with F denoting the SAR configuration. The parametrization now considers the effect of soil moisture (dielectric properties) and the effect of roughness (geometry of the surface) as independent. For the inversion of the model and vegetation water content estimation the vegetation contribution is neglected and only the attenuation of the signal returned by the underlying soil is considered, thus simplifying the term and by expressing it in dB units it leads to a linear

function of the surface variables:

$$\sigma^0 = \frac{-2 \times B \times w_c}{cos\theta + C(\theta) + D \times m_v} \tag{2.17}$$

with now σ^0 expressed in dB, θ the incidence angle [°radian], the coefficients B,C,D, w_c the total canopy water content and m_v the soil moisture. Results achieved when using this model, e.g., for LAI retrieval range between R^2 around 0.51 and 0.74 depending on the used SAR configuration.

Graham & Harris (2002) mention a limiting factor of the model, it‘s reliance upon multiple regression in parameter fitting. Each parameter has a physical meaning, which cannot easily be explained using regression techniques.
However, due to the need of regional and operational use it is more appropriate to employ easy to invert models with input parameters, which are easy to obtain without intensive ground truth.

2.2.4 Synergetic derivation of biophysical parameters using multi-sensoral remote sensing data

Since multispectral and microwave sensors respond to very different target characteristics, their role in crop monitoring can be seen as complementary (Liu et al., 2006). Both data types were already used successfully in previous studies for, e.g., classification, to monitor the crop condition or for yield forecasts. Whereby, in particular the all-weather capability of SAR and it‘s ability to fill time gaps in multispectral data was highlighted.
Brakke et al. (1981) analysed in their study data from a truck-mounted microwave SAR system and ground based spectral reflectance measurements individually and later in a combined way for extracting LAI and dry matter.
The combination was hereby achieved using a multiple linear regression with the scattering coefficient, canopy temperature and spectral reflectance ratios. They achieved the best results, when they correlated the LAI with the scattering coefficient and the normalized difference or the green/red ratio. For dry matter best results were achieved when using two frequencies and the green/red ratio. But on the other hand they also point out that their limited data set demands further research in that area.
In the study from Liu et al. (2006) it was tried to improve winter wheat yield estimation by using an Envisat ASAR image and a Landsat TM image. First they analysed by bivariate correlation the relationship to the single source data and then combined them in a multi-sensoral approach and thus improved their prediction accuracy.
Another option of a multi-sensoral approach was shown by Prévot et al. (2003) and Clevers & Van Leeuwen (1996) combining multispectral and SAR data in crop growth models. In the first study even though the simultaneous assimilation into a crop model was feasible the SAR information did not improve the results obtained when only using multispectral data. The study from Clevers & Van Leeuwen (1996) concluded, that when only a few multispectral data sets are available additional SAR data were able to improve the yield estimation. What confirms the main advantage of SAR lying in acquiring information on crop growth when other sensors fail.

2.3 Crop growth modelling

Crop growth modelling and final yield estimation of winter wheat represents another important issue within this thesis. Especially the question of assimilating remotely sensed information into an established model for improving the model performance is one major issue. In order to understand the whole process, firstly general assumptions about the phenologcial development of winter wheat will be made, then the used crop growth model CERES-Wheat will be introduced. Finally, different possibilities of assimilating external information into the model will be described with a main focus of the used re-initialization method.

2.3.1 Winter wheat and phenological development

The cereal addressed within this thesis is the common winter wheat (*Triticum aestivum L.*), which is the second most important food grain of the northern and southern temperate zones. In 2005 the world acreage of wheat was estimated roughly at 217 million hectare and a harvest of around 630 million tons. Average harvest was around 2900kg/ha and in Germany around 7500kg/ha.
The large harvest divergences and the global extent of cultivation implies a large variation in environmental and social settings. Which are e.g., sub-optimal agroecological conditions and situations prone to environmental degradation due to mismanagement (Jongschaap, 2006). Crop improvement and specific crop management have played an important role in agricultural research since 1970 for optimization of crop production. Whereby, especially the use of dynamic crop growth simulation modelling has played a substantial role.

In general the phenological development of winter wheat is dependent of the environmental conditions, thus it is hard to make a generalization of a standardized crop growth cycle. In Geisler (1970) the development of winter crops and the appropriate management practices are roughly allocated to the corresponding months and development stages (Fig. 2.14), which match the monitored growth cycles.
Generally, standard figures of sowing dates, seed rates [kg/ha] and target plant populations [plants/m^2] are applied (HGCA, 1998, 2000, 2002; McConnell, 1995). However, in practice the following local features are generally used to determine seed rate: time of sowing, soil type and seed bed quality and slug risk (HGCA, 2004). Drilling date is generally set by prevailing weather conditions and availability of labour. Weed control is required to keep weed populations at an acceptable level, usually described as the threshold value [plants/m^2]. However, the costs of applying the herbicide has to be below the value of yield lost, if the herbicide had not been applied. The agronomist must therefore plan for or identify the type of weed, e.g., black-grass, wild oats and cleavers, it´s density and it´s size (growth stage) in order to select the required type and dose of herbicide to use. Additionally, the agronomist must select a herbicide at a sufficient dose to control the weeds within the most infested patches, due to their non uniform distribution within the fields. Crop protection ensures an acceptable yield by protecting the crops against pests, diseases and lodging. For winter cereals variety selection and seed dressing are normally all the requirements to protect the crops over the winter period. Under certain weather conditions it might be necessary to spray for aphids preventing, e.g., gout fly and fruit fly especially for early sown crops. During spring winter cereal crops require fungicide

month	J	F	M	A	M	J	J	A	S	O	N	D
soil management (mechanical)			harrow					ploughing, seed bed				
chemical weed control												
fertilizing		N		N				K_2O; P_2O_5				
seed, harvest							harvest			seed		
phenological stage				stem elongation		flowering	ripening			germination	leaf development	

Figure 2.14: Winter wheat development and management practices, modified after: Geisler (1970).

applications, depending on feed or milling varieties. The type and dose rate of fungicide selection is based on factors like drilling date, crop variety, growth stage, canopy characteristics and disease pressure. Also spring growth regulators normally have to be applied to prevent crops from lodging. Whereby, large canopies and high plant population increase the risk of lodging. Crop nutrition supply ensures that crops have the sufficient nutrients (nitrogen, phosphorus, potassium, sulphur and other trace elements) to maintain healthy crop growth. Most significant and also difficult to measure is nitrogen. Thus, a soil index is used, based on the previous crop grown and soil type. Total amount of nitrogen is dependent on crop variety, soil type, soil nitrogen index and expected yield. Nitrogen is usually applied in three splits for wheat during spring. The amount of phosphorus and potassium a cereal crop requires, depends on the level of reserves in the soil, expected yield and whether the straw is removed (HGCA, 2004).

For identifying the different phenological development stages the BBCH-scale (*B*iologische *B*undesanstalt, *B*undessortenamt and *CH*emical industry) is used. It uses a decimal code system, which is divided into principal and secondary growth stages (Lancashire et al., 1991; Witzenberger et al., 41). Fig. 2.15 shows the development of winter wheat during a crop cycle and corresponding ground impressions by photographs.

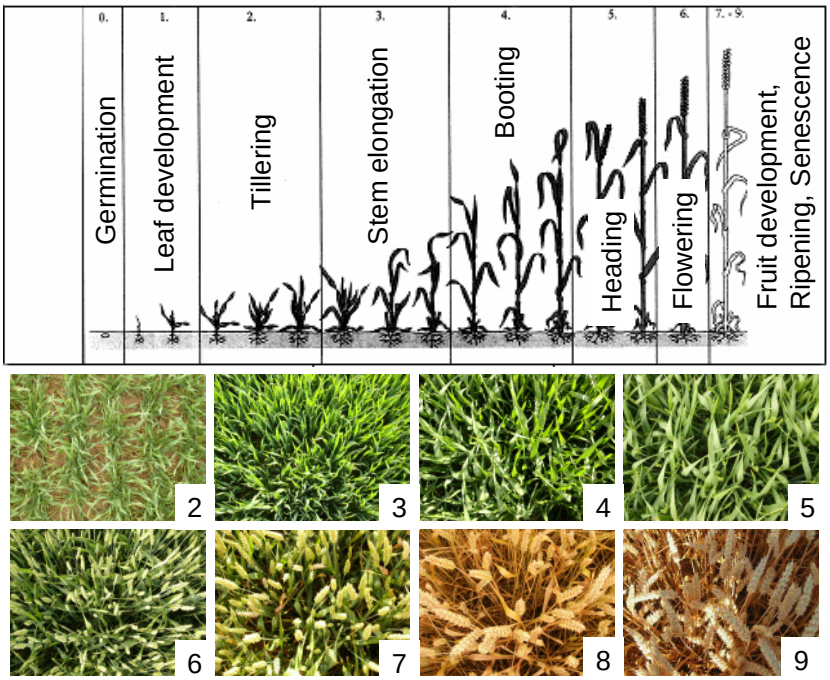

Figure 2.15: Phenological growth stages and BBCH-identification keys for cereals (schematic and photographs).

2.3.2 Introduction to different crop growth modelling approaches

Simulation models are developed to organize and quantitatively describe the processes and interactions in a systematic way, whereby all assumptions, inputs and results are quantified and thus can be made a subject of discussion. The basis therefore is the mathematical description of system processes linked through physical, physiological and biochemical laws. Especially, the increase in process knowledge and improvements in computer technology have led to the development of complex system models , e.g., crops-soil systems (Jongschaap, 2006).
Over recent years there has been a considerable progress in regional spatially explicit models of bio-geophysical processes of the land surface. The science of agroecosystem modelling started in 1960s with the later called *School of De Wit* (Wit, 1965). They developed increasingly complex dynamic simulation models, e.g., ELCROS (ELementary CROp Simulator model), BACROS (BAsic Crop growth Simulator model), or the simple universal model SUCROS (Simple and Universal CROp Simulator (model)), still used these days (Bouman, 1992). Since that time, many agroecosystem models have been developed with increasing complexity and multidisciplinary appendages. They model numerous vegetation types, different scales and a wide variety of applications.
Crop models can be grouped in various ways, whereby it is hard to categorize them due to, e.g., mixed forms. Normally one can divide them into empirical and functional (mechanistic) models. Empirical models emulate growth patterns on the basis of statistical principles from observed values. Hereby, they try to reproduce growth patterns under similar conditions.
The functional models, which are mainly of interest in this study, describe the physical, chemical and physiological process of the plant development. Therefore, they integrate measurable input parameters in defined time intervals and the development is mathematically defined (Jongschaap, 2006).
The current agroecosystem models, the mechanistic *S*oil-*V*egetation-*A*tmosphere (SVAT) models, incorporate not only biological and physiological knowledge of plants, but also try to integrate interactions between plants and their environment using functions quantitatively describing transfers over crop and soil boundaries and over crop internal boundary layers. In these models the vegetation state variables, e.g., development phase, dry mass, LAI are linked to driving variables, e.g., weather condition, nutrient availability and management variables. Output of theses models is usually final yield or accumulated biomass. They use computational iterations representing the time steps of the model, whereby at each iteration the vegetation state variables are updated based on input driving variables and values of state variables at the previous time step (Dorigo et al., 2007).

2.3.3 CERES-Wheat model

The DSSAT-software combines complex weather, management, soil, soil-plant-atmosphere, plant templates and plant models, as the CERES-Wheat (*Crop Environment Resource Synthesis*) model. In the next section therefore first a brief overview of the whole software system and the implemented models will be given, followed by a detailed description of the used plant module CERES-Wheat and an overview of some application examples.

2.3.3.1 Decision Support System for Agrotechnology Transfer (DSSAT)

The *D*ecision *S*upport *S*ystem for *A*grotechnology *T*ransfer (DSSAT) was developed by an international network of scientists, cooperating in the International Benchmark Sites Network for Agrotechnology Transfer project, for facilitating the application of crop models in a systematic approach to agronomic research (Jones et al., 2003). Knowledge about soil, climate, crops, and management for making better decisions about, e.g., transferring production technology from one location to another where soils and climate are different, are considered. The computer system aims at understanding how the environmental system and it‘s components function, for predicting the behaviour of the environmental system for given conditions. After modelling correctly the real world situation, experiments with, e.g., variable management or control systems can be performed. It was developed to operationalise the approach supporting decision-makers by reducing the time and human resources required for analysing complex alternative decisions and as a framework for researchers for integrating new knowledge and apply it to research questions. The main aims of the DSSAT are (Jones et al., 2003):

1. simulation of monocrop production systems considering weather, genetics, soil water, soil carbon and nitrogen, and management in single or multiple seasons and in crop rotations at any location

2. providing a platform for easy incorporating modules for other abiotic and biotic factors, e.g., soil phosphorus and plant diseases

3. providing a platform allowing easily to compare alternative modules for specific components to facilitate model improvement, evolution and documentation

4. providing a capability for easily introducing cropping system models into additional application programs

Fig. 2.16 shows the structure of the DSSAT software with a main driver (land unit model), and modules for the primary components (weather, soil, plant, soil-plant-atmosphere interface, management) making up a land unit in a cropping system (Jones et al., 2001). Collectively the primary components describe the changes in soil and plants occurring on a single land unit in response to weather and management. Each independent module has six operational steps, whereby the main program controls them. There are different types of application possible depending on the operational mode:

1. basic mode, providing interactive sensitivity analysis and comparison of simulated vs. observed field data

2. multi-year mode, allowing to evaluate the effects of uncertain future weather conditions on decisions made when all soil initial conditions are known

3. crop rotation mode, simulates crop rotation over a number of years

4. spatial-based mode, simulating crops over space (e.g., for precision agriculture or land use management)

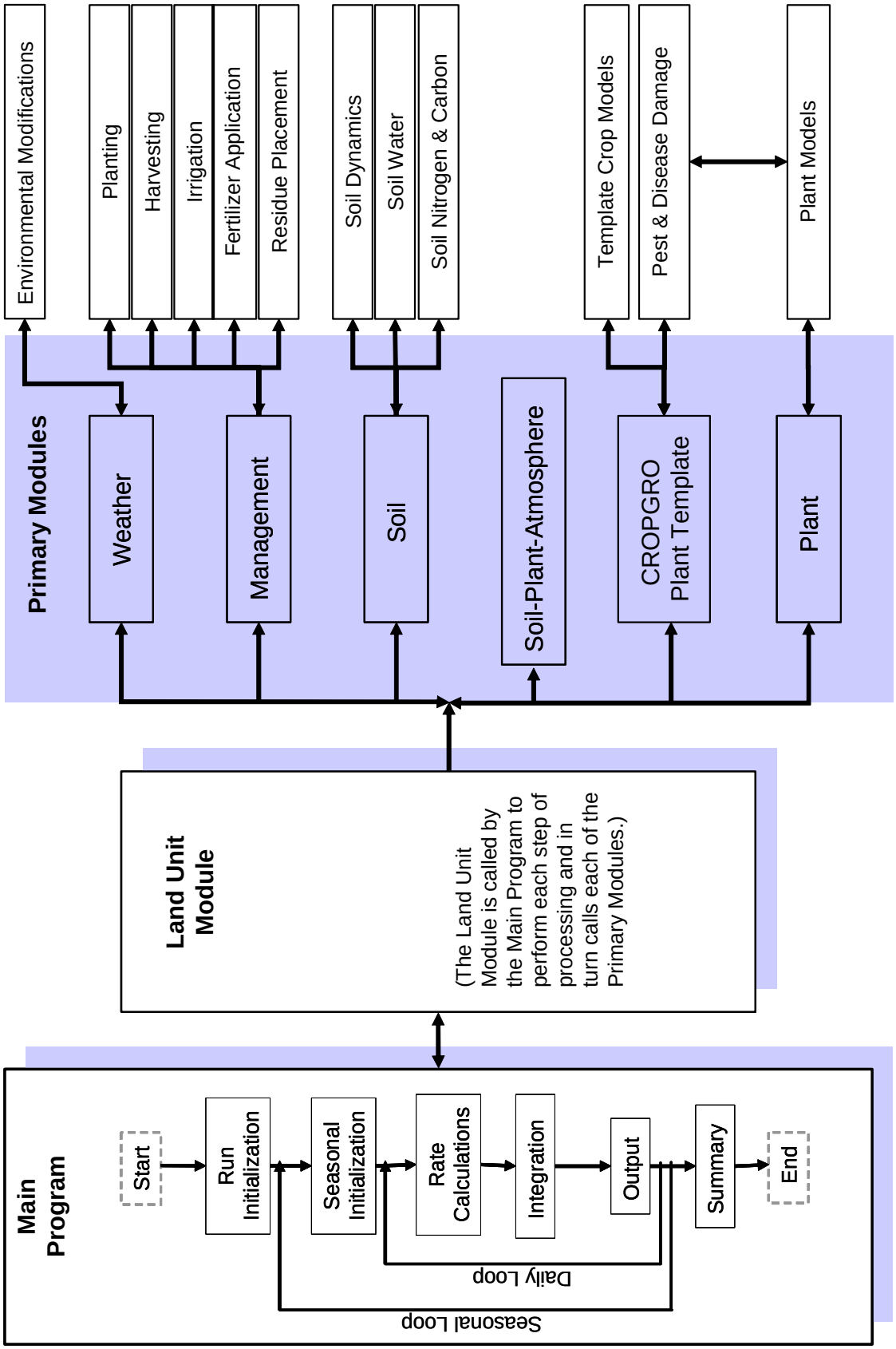

Figure 2.16: Overview of the components and modular structure of the DSSAT, modified after: Jones et al. (2003).

whereby, especially the last mode is of interest within the study. A more detailed description of the single components, as well as further references are given by Jones et al. (2003). In the next section only the plant module will be highlighted further.

2.3.3.2 Plant module: CERES-Wheat

In the plant module several different crop models are implemented, e.g., for maize, sorghum, barley, rice, potato, as well as the used CERES-Wheat model (Otter-Nacke et al., 1986). A general overview of the model and it´s inputs and outputs is given in Fig. 2.17.
For modelling, the plant life cycle is divided into 7 phenological stages (germination, emergence, terminal spikelet, end ear growth, beginning grain fill, maturity, harvest) (Jones et al., 2003). The rate of development is controlled by thermal time, or growing degree-days (GDD) computed based on daily maximum and minimum temperatures. The GDD govern the progress from one growth stage to another, which is influenced by cultivar-specific (genetic) coefficients listed in Tab. 2.3.

Daily plant growth is computed by converting daily intercepted photosynthetically active radiation into plant dry matter using a crop-specific radiation use efficiency parameter. Light interception is computed as a function of LAI, plant population, and row spacing. Daily growth may also be limited by water, nitrogen, temperature stress and it is sensitive to atmospheric CO_2 concentration. The above ground biomass has priority for carbohydrate, at the end of each day not used carbohydrate is allocated to the roots (Jones et al., 2003). Whereby, the roots must receive a specified stage-dependent minimum of daily carbohydrate for growth. The leaf area is converted into new leaf weight using empirical functions. During flowering kernel numbers per plant are calculated, based on the cultivar´s genetic potential, canopy weight, average rate of carbohydrate accumulation and stress factors, whereby the potential kernel number is user-defined. With the beginning of the grain filling period, the daily growth rate is calculated based on the user-defined potential kernel growth rate, temperature influences and assimilation availability of, e.g., carbon. The actual kernel growth is either until maturity or when the plant runs out of resources (Jones et al., 2003).

Table 2.3: Cultivar-specific coefficients (genetic) of the DSSAT CERES-Wheat model

coefficient	description	unit
P1D	Photoperiod sensitivity coefficient	% reduction/h near threshold
P1V	Vernalization sensitivity coefficient	%/d of unfulfilled vernalization
P5	Thermal time from the onset of linear fill to maturity	°C d
G1	Kernel number per unit stem + spike weight at anthesis	#/g
G2	Potential kernel growth rate	mg/(kernel d)
G3	Tiller death coefficient. Standard stem + spike weight when elongation ceases	g
PHINT	Thermal time between the appearance of leaf tips	°C d

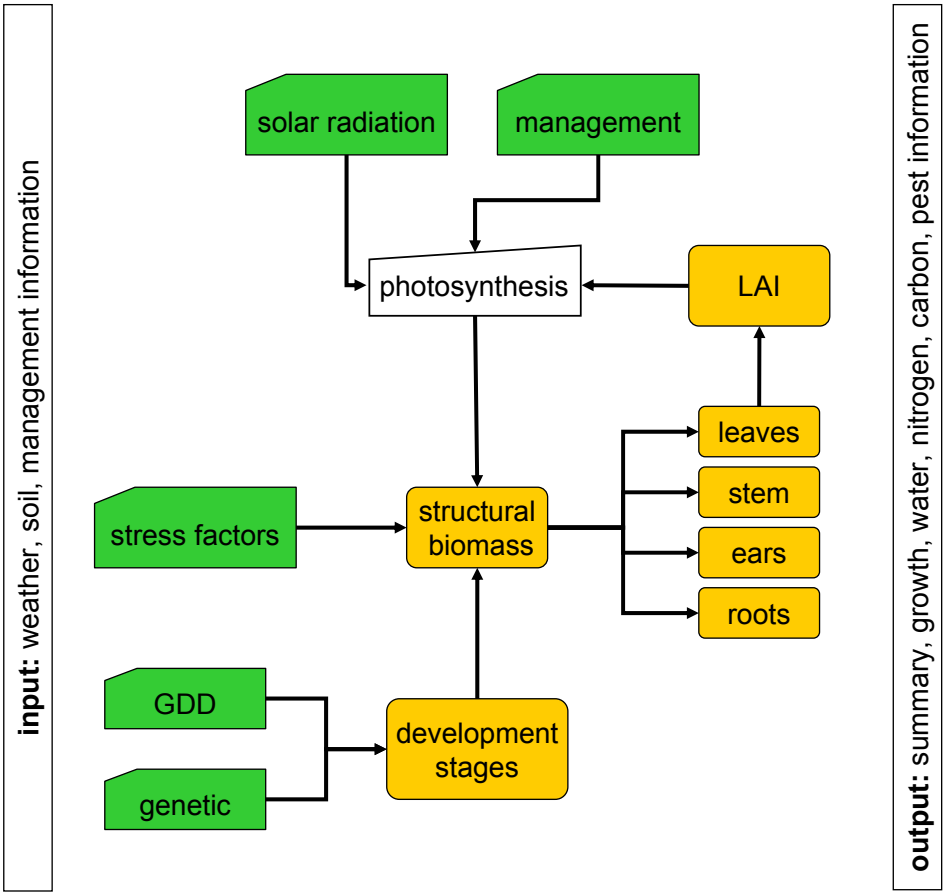

Figure 2.17: Overview of the CERES-Wheat model

In general for actually running the DSSAT models a minimum data set is needed for model operation. It encompasses data on the site where the model is to be operated, on the daily weather during the growing cycle, on the characteristics of the soil at the start of the growing cycle or crop sequence, and on the management of the crop, e.g., seeding rate, fertilizer applications, irrigations. Required weather data encompass daily records of total solar radiation, maximum and minimum air temperature and rainfall. If not all required weather data are available surrogate values are calculated by the use of climate statistics. Additionally, soil parameters for computing surface run off, evapotranspiration from the soil surface, and drainage are essential. Model evaluation requires the comparison of model outputs with real data. Whereby, these data should be an independent data set. Statements of adequacy of the model prediction include calculation of standard errors, root mean square error, and slope and intercept of regression of observed vs. predicted variables. For the model evaluation different data sets can be used, e.g., date of emergence, flowering, maturity, temporal LAI or canopy dry weight profile, canopy height or yield (Jones et al., 2003). The model outputs are a summary containing an overview of the main development events, water and nitrogen variables, yield and yield components. Also a growth output is produced with a summary of the growth balance variables (e.g., LAI, dry matter) over time. Further outputs are a water output, with the soil and plant water variables over time, nitrogen and carbon outputs, containing an overview of the soil and plant nitrogen and carbon variables over time, as well as a pest output, with the pest variables as a function of time.

Generally there are various types of applications of the DSSAT crop models for all countries (e.g., Asia, Europa, North America, Africa) and different types of application, e.g., crop management, fertilizer management, irrigation management, tillage management, variety evaluation, precision farming, environmental pollution, climate change and yield forecasting (Jones et al., 2003). The applications have been conducted by agricultural researchers using the model with field agronomic research and socioeconomic information to answer complex questions about production, economics, and the environment.
A main aspect of many studies was a consideration of weather influences on the performance of crops, interacting in complex ways with soil and management. Researchers have applied the model for studying uncertainty in crop production associated with weather variability and the associated economic risks that farmers face under such climate variability. Rosenzweig et al. (1995) analysed the impacts of climate change on agriculture. Therefore, they first evaluated model performance using data from current cropping systems and then used the models to assess the potential impacts of climate change on their cropping systems using different climate scenarios.
In Patel & Shekh (2005) the sensitivity of CERES-Wheat to weather and non-weather parameters was analysed. The results indicated that the model has the potential for assessing impact of climate change on the wheat production, as well as to management strategies, e.g., increased plant population density.
Another application issue is to use the models for advising farmers. Bannayan et al. (2003) have used the model successfully for yield forecasting within the growing season and showed that the model can aid in pre-season and within-season management decisions for cultural practices, such as fertilizer and irrigation applications and pest and disease management.
Another issue is the transferability of the model to other regions. Mavromatis et al. (2002) demonstrated the value of using routinely collected data from yield trials for estimating cultivar (genetic) characteristics and for evaluation for soybean. They used yield trial data from North America to show the robustness of model predictions across regions. The average simulated yields were within 2.5% of mean observed yields at each evaluated location.
Also precision farming has been an issue, for which the model can be integrated into a Geographic Information System (GIS) for diagnosing causes of yield variability and for prescribing variable management.
Braga (2000) evaluated the ability of the CERES-Maize for accurately describing the spatial variability on maize yields over two years for using it in precision agriculture research and decision support. He precisely measured soil water holding parameters, including initial conditions at planting and compared the simulated vs. observed maize yield. His study showed that the model reproduced observed grain yields for these conditions, when accurate soil, weather and cultivar (genetic) information were available.

2.3.4 Remote sensing data assimilation

Timely and accurate information on crop and soil status are essential information for the optimization of management decisions for crop production. Satellite remote sensing can supply information on plant status for large regions with high temporal resolution and can provide essential information for decision support. It allows accounting for spatial and temporal variations of environmental conditions, influencing crop growth and development, without extensive ground surveys. Therefore, combined approaches integrating remote sensing and dynamic crop growth models for regional yield prediction have been developed in several studies.

In Bouman (1992) the Cloud equations for SAR backscattering and an optical canopy radiation model were linked to the crop growth model SUCROS for sugar beet. The combined model was initialized and re-parametrized in order to fit simulated X-band SAR backscattering and/or multispectral reflectance values, to measured values. The simulated canopy biomass was compared with simulated canopy biomass using SUCROS with standard inputs and to ground truth information. The seasonal-average error was smaller using remote sensing information, than with SUCROS using standard input. The SAR backscattering data adjusted SURCOS only during early crop growth (exponential growth), whereas multispectral data still adjusted SUCROS until late in the growing season.
In Dente et al. (2004) the information retrieved from a multi-temporal series of C-band HH/VV backscatter ratios were assimilated into the CERES-Wheat crop model. A sensitivity analysis showed that inaccurate knowledge of some model inputs, concerning soil properties and crop management, can lead to erroneous predictions. When adopting a re-initialization assimilation strategy, significant improvements in the model estimations were obtained.
The primary objective in Doraiswamy et al. (2003) was the evaluation of integrating parameters retrieved from satellite data in a crop model to simulate spring wheat yields at sub-county and county levels. The input parameters extracted from remote sensing data provided spatial integrity, as well as a real-time calibration of model simulated parameters during the growing season. Ensuring that the modelled and observed conditions are in agreement.
In Clevers & Van Leeuwen (1996) multispectral and microwave remote sensing data were used in combination for crop growth monitoring. Whereby, especially the synergistic effect of using both data types for estimating LAI was analysed by studying different data acquisition scenarios. Models were then inverted to obtain LAI estimates during the growing season for calibrating the crop growth model to actual growing conditions. The findings were that simultaneous observations did not improve estimates of LAI over multispectral data alone. Thus, only when few multispectral recording dates are available SAR recordings give a slight improvement of the results of crop monitoring and yield estimation compared with the multispectral data alone. In the absence of multispectral data, SAR data yielded a significant improvement in yield estimation compared with the case of no remotely observed information. Confirming the main advantage of SAR lying in acquiring information on crop growth when other techniques fail.
Prevot et al. (1993) coupled the STICS crop model for wheat with radiative transfer (RT) models in the solar and microwave domain. Permitting the simulation of the temporal variation of remotely-sensed data over wheat canopies. They recalibrated some of the STICS parameter constraining the evolution of the simulated canopy variables. Also a sensitivity analysis of the coupled STICS+RT model was carried out selecting the parameters of STICS with the largest influence on remotely-sensed observables and on some canopy variables (dry biomass and LAI). The sensitivity analysis pointed out the parameters, which should be recalibrated in the assimilation process to adequately simulate the time course of the remotely-sensed data. In a second step they also compared the benefits of assimilating multispectral data alone, microwave data alone and both data together into STICS+RT. However, the introduction of microwave data did not improve the accuracy of the simulated values, thus, simultaneous assimilation is only of great interest when cloud cover limits the number of available multispectral images.
Jongschaap & Schouten (2005) used multispectral remote sensing data to estimate winter wheat areas in the South-east of France. Microwave data were then further used to estimate regional wheat heading/flowering dates for calibrating a wheat growth simulation model calculating wheat yields, subsequently estimating regional wheat production. The results demonstrated that findings from point-based simulation models can be applied at spatially higher levels with the aid of remote sensing data.

In Dente et al. (2007) leaf area index retrieved from ENVISAT ASAR and MERIS data were assimilated also into the CERES-Wheat model. Whereby, the objective was to improve the wheat yield predictions at catchment scale. The assimilation method consisted in re-initializating the model with optimal input parameters, allowing a better temporal agreement between the LAI simulated by the model and the LAI estimated by the remote sensing data. The study shows that the lack of MERIS data during critical phenological stages can be overcome by integrating SAR data. The combined use of all LAI maps retrieved from both data types leads to an error equal to the case of only using MERIS data, without time gaps during critical phenological stages.

The term assimilation, which will be used here, originally comes from the Latin word *assimilatio*: *alike making*. It is used mainly in the context of metabolism physiology, describing the process of producing body own substances from body extraneous substances by energy use (Leser, 1998). But it is also used in other disciplines like ethnology, describing the process of the adjustment of people into another ethnic group, taking over their behaviour and system of values. Or in the context of geology, describing the process of fusion from foreign stones with magma. In the framework of this work the objective of data assimilation is to characterize the state of an agroecosystem by combining information from various sources such as mathematical models, remote sensing data or time of the variables of interest. Within a data assimilation framework it is necessary to distinguish (1) driving variables forcing the system (e.g., net primary production, evapotranspiration), (2) state variables providing a complete description of system behaviour (e.g., LAI, dry matter), (3) model parameters characterizing the relationship between driving variables and state variables (e.g., development or growth parameters), and (4) output variables, which are observable functions of the state variables (e.g., final yield) (Dorigo et al., 2007). The coupling between a crop model and satellite data can be achieved by the use of different methods (Bach & Mauser, 2003; Delecolle et al., 1992; Dorigo et al., 2007; Maas, 1988a; Moulin et al., 1998).

By *the direct use of a driving variable*, which is estimated from remote sensing information in the model, one assumes that accurate remote sensing information is available at adequate time steps, e.g., daily or weekly. The model does not determine values of that variable. Due to cloud coverage and the repetition rate of the satellites this is rarely the case. To overcome the problem of this *forcing method*, gaps may be filled by some interpolation procedure, usually by fitting an empirical curve of time evolution of the state variable to the remote sensing data. The most popular approach of this method is the already mentioned *efficiency equation* of Monteith (1977). He tried to overcome the problem of remote sensing data availability by *updating a state variable*, e.g., LAI derived from remote sensing data to obtain a simulation in agreement with the variable derived from the observations. The remotely sensed observations are considered as an absolute reference for the model simulation. This assumes that the formally adequate model may be inaccurately calibrated. Studies, e.g., by Clevers & Van Leeuwen (1996) and Maas (1988b) show that the estimation of the biomass time profile and final yield are improved as compared to results using the forcing method.

The *re-initialization* of the model, e.g., the adjustment of an initial condition defined in the management input for obtaining a simulation in agreement with the remotely sensed observations is used in this study. Hereby, e.g., initial conditions as sowing date or wilting point are adjusted until the error between the temporal remote sensing series and the model prediction is minimized. Fig. 2.18 illustrates the general proceeding using this method for re-tuning an initial condition.

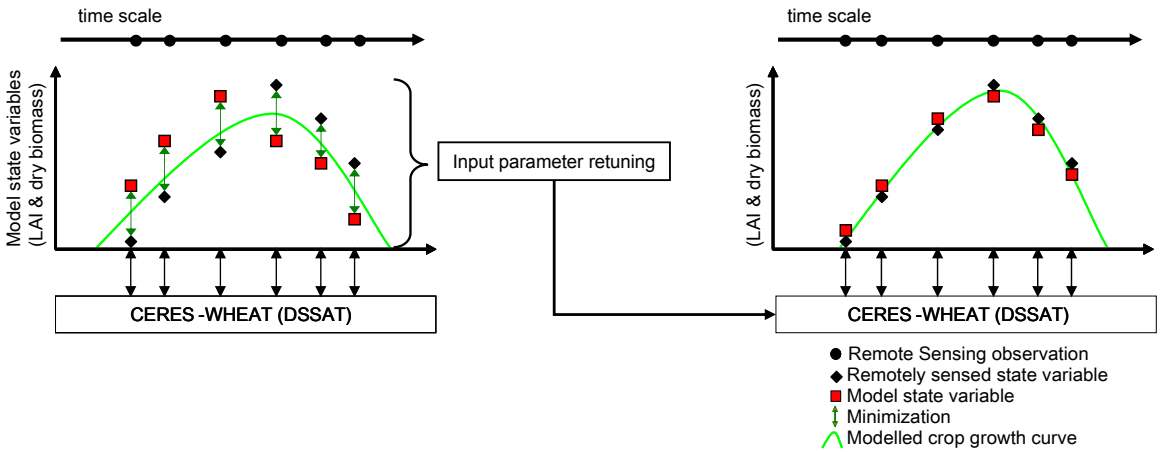

Figure 2.18: Remote sensing data assimilation by re-tuning the initial conditions, modified after: Delecolle et al. (1992).

In Fig. 2.19 the exact interaction between the remotely sensed informations and the CERES-Wheat model is displayed. By comparing the modelled variables with the observed information, derived from satellite data the input parameter management is retuned. In order to identify the model input parameter mainly affecting the outputs a sensitivity analysis has to be carried out taking into account management and e.g. soil characteristics. Afterwards the simulation is repeated and the temporal profiles and the statistical parameters , e.g., standard deviation, mean are compared again.

The merits and demerits of data assimilation strategies are that the observed state variables derived trough remote sensing at regional scale also might contain some observation errors (Bastiaanssen, 1998). In the case of the *forcing* method, the model rejects own information, and follows the observed state variables, including the observation errors. Thus, the error in the remotely sensed state variables is propagated into the model. The *re-initialization* and *updating* methods have more flexibility in the assimilation of remotely sensed state variables and their associated errors into the model. As the physical description of the underlying process is an acceptable representation of the natural system, the *re-initialization* method is expected to give more representative input parameters and improve model predictions. This only applies if there were sufficient observations, and the observation error is small. A disadvantages of this method is the large amount of computing time, due to the optimization procedure. Adopting the *updating* method requires only one run of the model, thus significantly reducing the computing time. However, the drawback of this approach is the error propagation of the measured variables into the model state variables of the system. Moreover, for the *updating* method it is necessary to adjust model state variables during the model run, which often intervenes deeply into the model structure and processing loops.

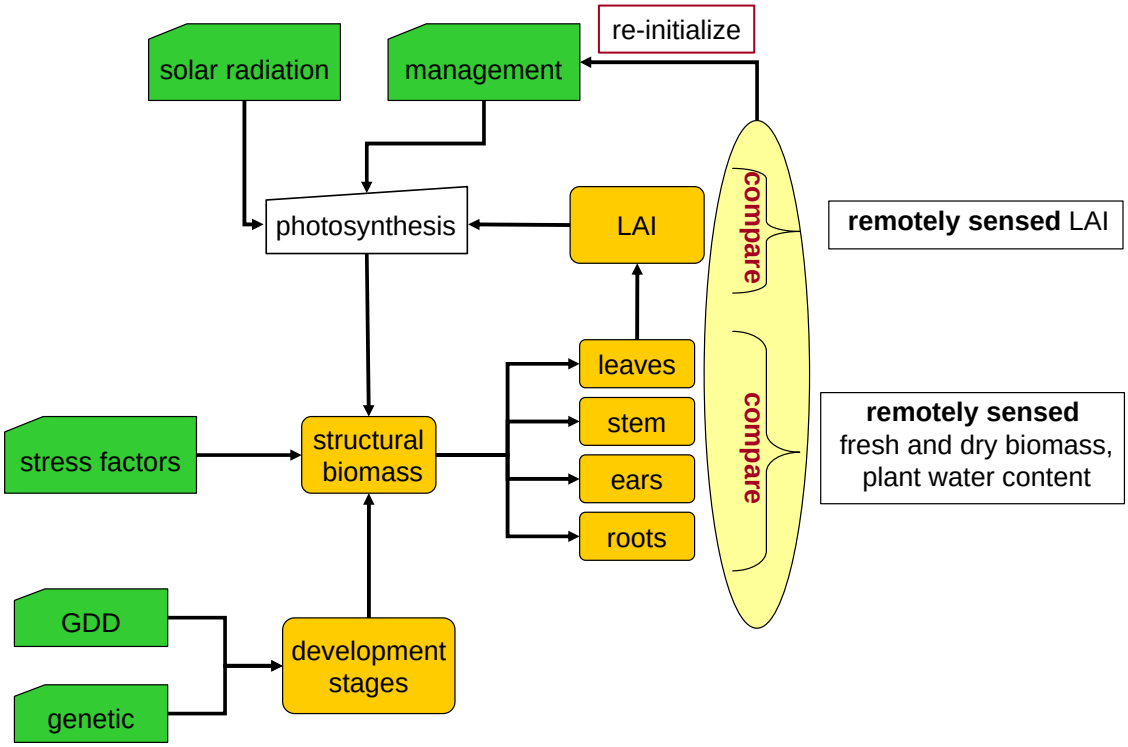

Figure 2.19: The CERES-Wheat model and the linkage with the remotely sensed information.

3 DATA

In the following chapter firstly all the necessary ground truth data will be described and the research area. The sampling design will be exemplified and all necessary additional processing steps and considerations. In a next section remote sensing sensors used within the study will be described. Whereby, also the respective preprocessing steps of the different used remote sensing data will be summarized. A flowchart (Fig. 3.14) gives a rough overview of the individual processing steps of the multispectral and SAR data.

3.1 Ground truth

3.1.1 Region Meckenheim and the research farm Klein-Altendorf

The research area, a region near the cities of Bonn and Cologne in North Rhine-Westphalia (Germany) strechtes from the Köln-Aachener bay on the *Mittelterrasse* of the river Rhine, 62m asL, to the Rhine-Sieg area close to the town Meckenheim, 175m asL (Fig. 3.1). The climate is maritime with mild summers and winters. Average yearly temperature is 9.2°C, average rainfall 600mm with 1534 sun hours per year and a vegetation period of 165-170 days starting from mid March until beginning of December. General information about the soil characteristics in the region were given by the geological survey North Rhine Westphalia (Dworschak et al., 2001). They divided the soils by soil type, soil horizon, ground water and stagnant water. More detail about the soils of the research farm were provided by Lehmann & Pätzold (1996), they also analysed, e.g., pH, organic content or the particle diameter. Main soils are *Luvisols* (FAO system) from loess within the whole research region. Main farming crops are root crops, winter wheat, winter barley, rape and arable crops.
The research farm Klein-Altendorf (Fig. 3.1, longitude: 6°59; latitude: 50°37), where the ground truth campaigns were carried out, belongs to the University of Bonn and is located in Klein-Altendorf close to Meckenheim. It‘s topography is relatively flat with an inclination of 0.5-1°. The groundwater level lies at 20m. The monitored field sizes range from 6.9 up to 8.5ha.

3.1.2 Field experiments 2005 and 2006

Within the years 2005 and 2006 two and respectively four winter wheat fields were monitored in weekly intervals starting in April (stem elongation stage) until harvest at the beginning of August (Fig. 3.1). The fields were treated normally in respect to, e.g., applications, seeding and

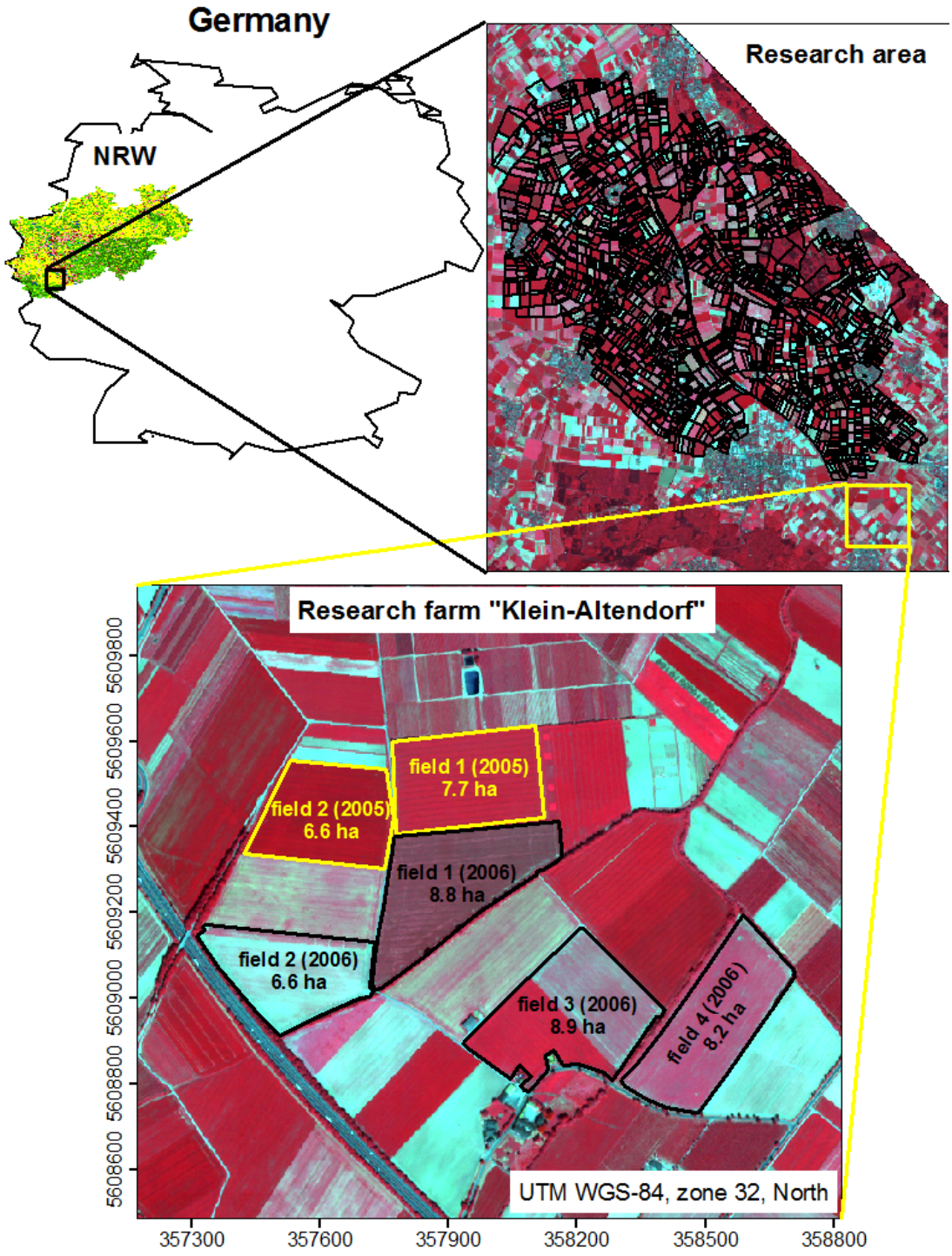

Figure 3.1: North Rhine Westphalia (NRW) in Germany with the research region and research farm *Klein-Altendorf*.
(Source: Heinzel and NRWPro Research project, *ZFL*)

harvest. The collected ground data included field properties (i.e. row spacing, row direction), wheat phenological stage, plant population density, wheat biomass, including water content, fresh and dry weight of winter wheat separated into stems, leaves and ears, canopy structure (e.g., height, leave and ear geometry), canopy coverage and profile, leaf area index (LAI), volumetric soil moisture and soil texture. The field measurements were divided into non-destructive variables (i.e. height) and destructive variables (e.g., biomass) each with separate sampling points or areas.
In general the measurement points were chosen considering the properties of the respective field (e.g., field size and soil properties), the resolution of the used remote sensing data and statistical considerations. Because of the low resolution from, e.g., Landsat 5 TM or ERS-2 and the speckle problem of the SAR images, the sampling results for each field were averaged. The amount of sampling points ranges from three to eight destructive and non-destructive measurements. They were all located on the boundaries of the fields, being far enough from the borders to be undisturbed, but also leaving large undisturbed areas in the middle of the fields (Fig. 3.2).

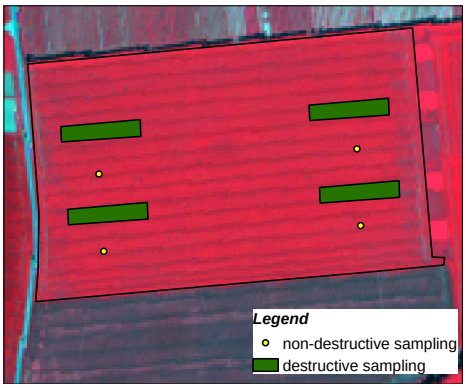

Figure 3.2: Sampling design exemplarily for winter wheat field 1 (2005) with a false colour QuickBird imagery (04/22/2005) as background.

Destructive sampling

Fig. 3.3 illustrates the biomass sampling design exemplarily for winter wheat field 1 (2005). Destructive biomass sampling plots for each date were $4.5m \times 5m$, whereby the biomass cutting was done in the first part ($1.5m \times 2m$) of each plot. The leftover remainder area to the next sampling area guaranteed that previous sampling did not influence the new sampling on the next date. The cutting volume for each sample cut was three examples from different plant rows over a length of 0.5m, which ensured getting representative samples. This was done twice for each cutting area in order to get a representative average over the whole field. The numbers (1-16) in the sampling plots indicate the different dates of the ground truth campaigns.
Samples were separated into stems, leaves and ears, depending on the phenological stage, and then weighed to obtain the fresh biomass. Plant water content (PWC) and dry biomass (DM) were also collected separately for each plant component by drying the samples in an oven at 105°C for 24 to 48h, depending on the phenological stage. This was done due to the fact that the ears need longer for drying when they are in the corn filling phase then just before harvest. The values of the measured biophysical parameters were all calculated into kg/m^2, due to comparability reasons and retrieval from the remote sensing data. For the later assimilation process into

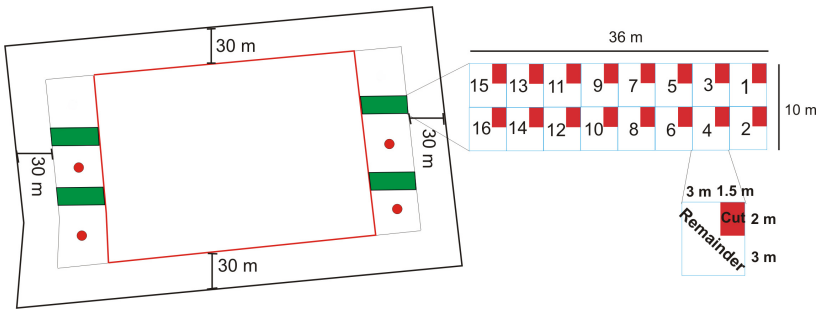

Figure 3.3: Layout of destructive biomass sampling, exemplarily for winter wheat field 1 (2005).

the CERES-Wheat model, they were converted into kg/ha, whereby one ha equals 10,000m^2. Fig. 3.4 shows the total fresh and dry biomass development in kg/m^2 averaged over all fields monitored within that specific year. The generalized growth curves of the biomass development from the two years studied are rather similar. While the fresh biomass curves differ a little from each other, the dry biomass curves run nearly parallel to each other. Only end of May and end of June the 2006 dry matter curve shows a slight decrease, probably due to measurement errors in the laboratory while drying, because this trend cannot be clearly observed in the fresh biomass curves. In order to get an impression of the homogeneity of the monitored fields and to ensure that the averaged field measurements represent the field condition correctly, a statistical analysis of the individual measurements within each field was done. Therefore, arithmetic average, standard deviation, and the percental standard deviation of the single field measurements were analysed. Tab 3.1 illustrates exemplarily the results of the destructive sampling in 2005. By the interpretation of the percental standard deviation of each field it becomes clear that the fields were rather homogeneous, only for one measurement date it exceeded 10%. For the 2006 data the results were comparable.

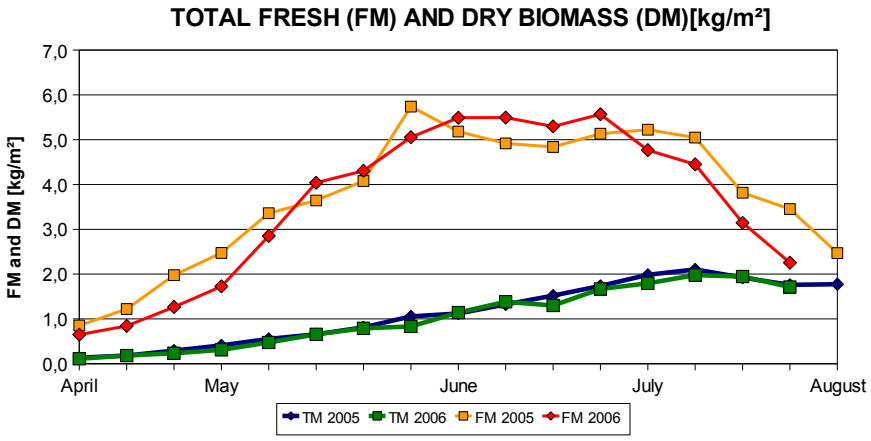

Figure 3.4: Fresh and dry biomass curves [kg/m^2], averaged over all fields.

Table 3.1: Statistics of the destructive sampling 2005, exemplarily for dry biomass [kg/m^2]. With the arithmetic average [kg/m^2], standard deviation, percental standard deviation and variance for field 1 and field 2 of the field campaign 2005.

	field 1			field 2		
date	$\overline{a}$	s	$spercent$	$\overline{a}$	s	$spercent$
04/11/05	0.09	0.02	6.3	0.18	0.03	7.5
04/18/05	0.14	0.02	4.0	0.23	0.04	9.1
04/25/05	0.23	0.04	5.8	0.36	0.08	11.3
05/02/05	0.35	0.02	2.1	0.47	0.05	5.1
05/09/05	0.49	0.09	6.0	0.61	0.05	4.3
05/17/05	0.56	0.09	6.0	0.75	0.1	6.9
05/23/05	0.72	0.06	2.8	0.9	0.14	7.46
05/30/05	1.0	0.11	3.7	1.11	0.17	7.87
06/06/05	1.08	0.15	4.9	1.16	0.12	5.29
06/13/05	1.27	0.15	4.2	1.38	0.09	3.38
06/20/05	1.47	0.19	4.6	1.56	0.22	7.09
06/27/05	1.48	0.19	4.7	1.99	0.12	3.00
07/04/05	2.01	0.4	7.1	1.96	0.25	6.42
07/11/05	1.97	0.2	3.7	2.23	0.35	7.83
07/18/05	1.89	0.23	4.3	1.94	0.18	4.84
07/25/05	1.72	0.2	4.2	1.8	0.39	10.91
08/01/05	1.59	0.16	3.6	1.96	0.2	5.05

Non-destructive sampling

Tab. 3.2 displays all non-destructive parameters divided into parameters measured in weekly intervals and constant parameters.

Constant parameters
Information about the exact seed dates, species, field sizes, previous crops were given by the local farmer Mr. Huober, who manages the research farm (see Tab. 3.3). Row direction and spacing were recorded once each vegetation period for the separate fields. The plant population density (ppd) [$plants/m^2$] was assessed during an early vegetation stage, when the different plants could still be distinguished from each other. Counts were done 4 times on each field for different rows over a length of 1m. Harvesting was between July to August by the local farmer. The actual grain yield was recorded by an on-board combine harvester, whereby measurements were taken every 3-5m, depending on the actual speed of the tractor. These point measurements were then interpolated using an ordinary kriging method, in order to produce actual final yield maps. Fig. 3.5 displays these grain yield maps for the different fields monitored in the years 2005 and 2006, transferred into kg/ha for comparability reasons. The coding of the fields stands for: WW: winter wheat, numbers from 1-4 the respective monitored fields and -05 indicating the studied year. For example: WW1-05 codes the winter wheat field 1 monitored in 2005. Regions with very low grain yield (red) could be due to headland, especially at the boarders of the fields, or to animal damage and differences in the soil properties, e.g., because of old roads or old construction sides. Due to the resolution of the remote sensing sensors grain yield was averaged for each field.

Table 3.2: Non-destructive parameters measured

constant parameters	weekly measured parameters
seed date	phenological stage
row direction and spacing	plant height
plant population density	plant geometry (leave and ear)
harvest	canopy coverage
soil parameters	leaf area index (LAI)
	vegetation profile
	volumetric soil moisture
	soil roughness
	agricultural management

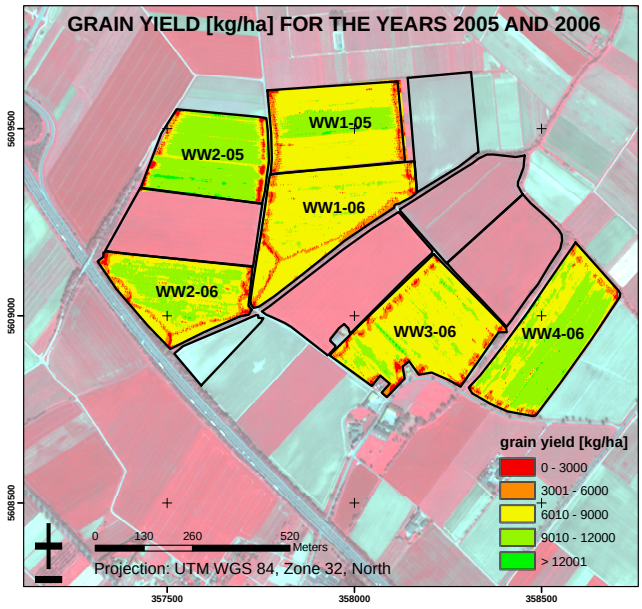

Figure 3.5: Grain yield [kg/ha] for the winter wheat fields monitored in 2005 and 2006.

Table 3.3: Constant field parameters during the research years 2005 and 2006

field name	seed date	species	field size ([ha]	row direction [°]	row spacing [cm]	ppd [plants/ m^2]	previous crop	grain yield [kg/ha]	harvest date
WW1-05	11/02/2004	Dekan	7.7	90	12.5	232	s.b.	7970	08/11/2005
WW2-05	10/16/2004	Tommi	6.9	90	12.5	190	s.b.	9850	08/10/2005
WW1-06	10/13/2005	Tommi	8.8	70	12.5	170	w.r.	7440	07/26/2006
WW2-06	11/14/2005	Tommi	6.6	90	12.5	180	s.b.	7820	08/07/2006
WW3-06	10/07/2005	Tommi	8.9	160	12.5	160	s.b.	6980	07/25/2006
WW4-06	10/11/2005	Dekan	8.2	220	12.5	220	s.b.	8920	07/27/2006

s.b.=sugar beet; w.r.=winter rye

Weekly measured parameters
The sampling design for the non-destructive parameters is also displayed in Fig. 3.2, whereby the dots indicate the single measurement points.

Phenological stage The phenological stages of each field were recorded over the whole measurement campaigns and documented in the BBCH code (Lancashire et al., 1991). Fig. 3.6 illustrates the phenological development of two wheat fields in the years 2005 and 2006. What becomes obvious is that the development until June (*heading - flowering*) is nearly parallel. After this stage the development for the crops in 2006 becomes a little faster, because of better meteorological conditions. Another point of interest is that harvest was rather late in 2005, due to a rainy period.

Plant height The plant height was measured five times at each of the non-destructive measurement points and was then averaged over the whole field. Plant height was also a good indicator for wind blow within the fields, which was a problem especially in 2005.

Plant geometry At each measurement point the geometry of three plants was measured exemplarily. Geometry features were the leaf amount, leaf length, leaf width and the percentage of healthy green leaf. The information was needed to model the leaf area index in a non destructive way in order to compare it with the performances of the Plant Canopy Imager CI-110.

Canopy coverage Also measured was the canopy coverage in percentage. Which was estimated on field using an estimation frame ($50 \times 50cm$ with a $10 \times 10cm$ raster). The other estimation method was by taking a nadir digital picture at two points per sampling point. These were then classified into the classes: shadow, vegetation and soil using a Maximum Likelihood classification procedure. When comparing these two methods, it becomes obvious that field estimations are mostly overestimated during rather early phenological stages and just before harvest (by 10-25%). During the other phenological stages differences between them only added up to around 5% (Fig. 3.7).

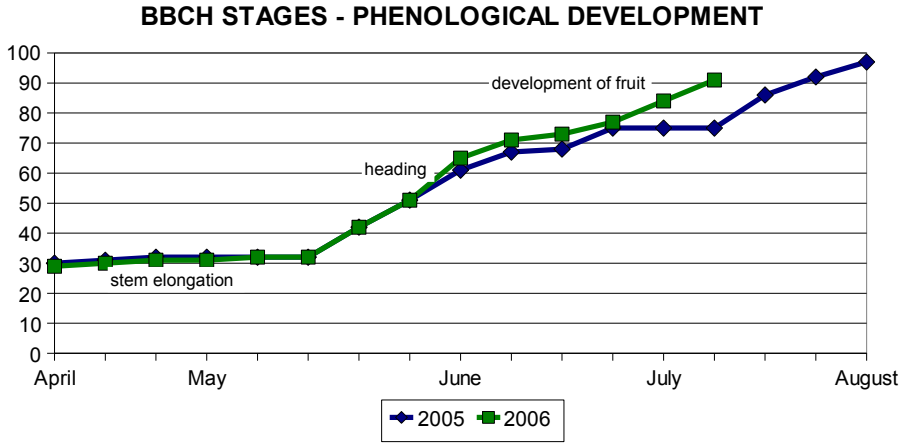

Figure 3.6: Phenological development of two exemplarily wheat fields (2005, 2006) coded in the BBCH system.

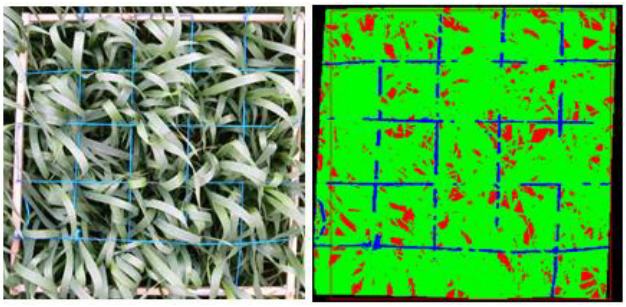

Figure 3.7: Digital picture (05/09/2005) of the wheat coverage with the estimation frame and the classification result of the image.

Field coverage estimation on field was 90% the classification estimated 85%.

Leaf area index The leaf area index (LAI) is one of the most important parameters for vegetation canopy characterization and is used by many scientists for different purposes. It was defined by Watson (1947) as the total one-sided area of leaf tissue per unit ground surface area. LAI is one state parameter of many crop growth models, whereby its values directly contribute to the determination of the potential plant assimilated absorption (Guissard et al., 2005). Therefore, it is derived from remotely sensed vegetation indices, which are cross-calibrated on ground-based LAI estimates. Thus, it is important to determine it correctly in the ground truth campaign, wherefore different destructive and non-destructive methods were tested during the field campaigns. In Fig. 3.8 (A) the destructive measurement method is illustrated, for which twice during the vegetation periods several cuts of $50 \times 50cm$ were taken within each field. The leafs and steams were then separated and fixed to a white background with a defined area ($50 \times 50cm$), whereby around three to four sheets were necessary for each sampling point. Sheets were then photographed with a nadir viewing geometry. These images were afterwards classified using a Maximum Likelihood classification algorithm into the classes background and leaf. Over the relative area of the sheet covered by the class leaf and the size of the sheet the actual LAI was then calculated. Problematic with this method is the destructive character and the expenditure of time, especially when a high spatial and temporal resolution is necessary.
Also illustrated in Fig. 3.8 (B) is an image captured by the digital Plant Canopy Imager CI-110

using a fish-eye lens (150). A special software (CI-110) calculates the solar beam transmission coefficients, using a user-defined number of zenith and azimuthal divisions. For the calculation of the LAI, diffuse radiation transmission coefficients, mean foliage inclination angles and plant canopy extinction coefficients the gap-fraction inversion procedure is used (Norman & Campbell, 1989; Colombo et al., 2003):

$$LAI = \frac{ln(P(\theta))cos(\theta)}{G(\theta)} \tag{3.1}$$

with $P(\theta)$ as the gap fraction and $G(\theta)$ as the G-function independent of the leaf-angle distribution. Problematic is that actually the plant area index (PAI) is calculated, because, e.g., stems are still included. This indirect measurement method represents rather effective LAI for agricultural crops, but no published results of cross-comparison are available yet (Bréda, 2003). Measurements were taken twice per non-destructive sampling point and were then averaged over the whole field.
In Fig. 3.8 (C) another non-destructive LAI estimation method used is displayed. It is based on leaf dimensions such as length, width and a conversion factor (Hoyt & Bradfield, 1962; Tewolde et al., 2005):

$$LAI = C \times \left(\frac{\sum(l \times w)}{n} \times ppd \right) \tag{3.2}$$

whereby, C equals the conversion coefficient, l the leaf length, w the leaf width measured at the broadest location and ppd the plant population density in plants per m^2. The value of the conversion factor was estimated from the field data (destructive LAI measurement) using regression analysis.
Fig. 3.9 displays the results of the different LAI measurement methods and a modelled LAI curve as reference for the year 2006. The modelling was done with the DSSAT software and the appropriate weather data for 2006 (section 2.3). When comparing the different measurement methods with the modelled curve, it becomes obvious that the measurement results are all lower than the modelled curve and never exceed a LAI of 2.5, whereby winter wheat can have LAIs around 5 or even higher. Especially, the leaf dimension method produces rather low LAIs, but one should take hereby into account that for this index only green leafs were taken into account, while with the CI-110 canopy Imager brown leafs and the stems are also regarded as vital leaf tissue. The destructive method gives rather reasonable results, but measurements were only done during vegetation stages with relatively low LAI. This also effects the leaf dimension method, because the conversion coefficient is calibrated with these data. Thus, all these different measurement methods contain different error sources. Further calibration of the retrieval models for the remotely sensed data is done with the LAI captured by the leaf dimension method. For these method the highest correlations to the remote sensing data were found and the data series using this method is the densest.

Vegetation profile The vegetation profile was recorded in weekly intervals for the years 2005 and 2006 and is a good indicator for the roughness of the vegetation and the height differences within the crops. For the assessment, digital pictures of a board, vertically in placed the crops, were taken from a horizontal viewing point. The board was $100 \times 80cm$ (height $\times$ width) with a $10 \times 10cm$ raster. In a first processing step digital pictures were rectified with the assistance of the screened board to UTM WGS-84, zone 1, north projection with a pixel size of 0.1m. In a second step the vegetation profiles (see picture Fig. 3.10) were digitized using a point shape with 2cm sampling raster. The Y-coordinates of each point were than finally used for

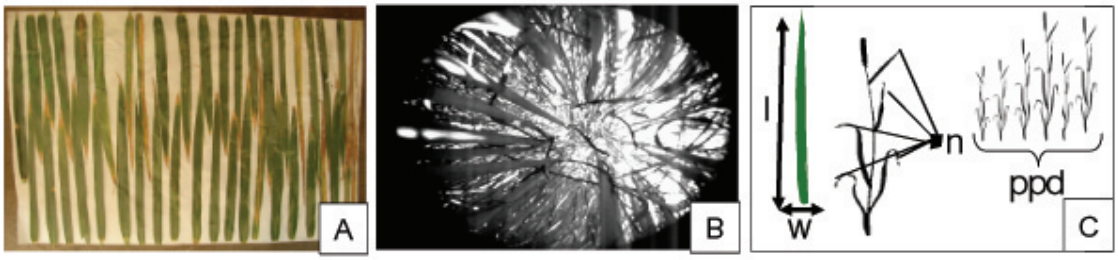

Figure 3.8: Different methods for LAI estimation.

A= destructive measurement; B= LAI CI-110 measurement; C= leaf dimension measurement

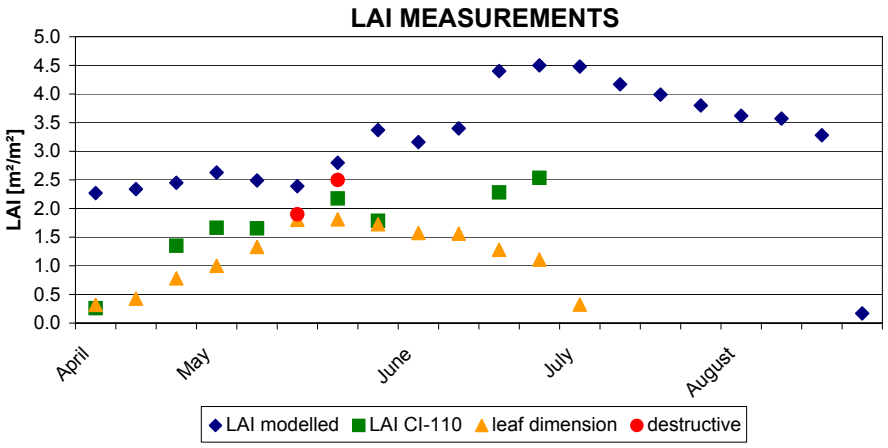

Figure 3.9: LAI measurements from the different techniques used.

the calculation of the standard deviation of vegetation height (s) and the correlation length (l). Both are presented in units of centimetres. For the series z_i; l=1,2,...N s is given for the discrete one-dimensional case by:

$$s = \left[\frac{1}{N-1} \left(\sum_{i=1}^{N} (z_i)^2 - N (\overline{z})^2 \right) \right]^{\frac{1}{2}} \tag{3.3}$$

where,

$$\overline{z} = \frac{1}{N} \sum_{i=1}^{N} z_i \tag{3.4}$$

and N is the number of samples. The relationship between height above an arbitrary plane of one point located at a point a and the height of another point a_ω distant from a can be statistically expressed in form of an autocorrelation coefficient. The variation in the value of the autocorrelation coefficient as the distance between the two points increases is referred to as the autocorrelation function. On the occasion this function can be mapped to a mathematical function such as the exponential or Gaussian (normal) distributions. The correlation length (l) is the displacement from the original point, a, when there exists no statistical relationship between two points. The normalized autocorrelation function, $r(a_\omega)$ in the discrete case is given by:

$$\rho(a^{'}) = \frac{\sum_{i=1}^{N+1-j} Z_i Z_{j+i-1}}{\sum_{i-1}^{N} Z_i^2} \tag{3.5}$$

for a spatial displacement $(a_\omega) = (j-1)Dx$, where j is an integer. The surface correlation length l is usually defined as the displacement (a_ω) for which $r(a_\omega)$ is equal to $1/e$:

$$\rho(a^{'}) = 1/e \tag{3.6}$$

Considering, e.g., an agricultural field and a flat asphalt surface. The randomly distributed, non-uniform shape and size of the agricultural field should yield relatively short correlation lengths, because of a quickly diminishing relationship between displacement from the origin and the surface height measurement. Alternatively on a flat, asphalt surface with almost no variations in surface height, even points with a large displacement from the origin will be highly correlated with the height measurement of the origin point.

Both of these variables were then averaged for each field, Fig.3.11 displays the temporal development exemplarily for winter wheat field 1 (2006) for the dates being later correlated with remote sensing data. The temporal trend, especially of s is a steady linear increase until the end of blossom (anthesis)(06/14/2006). After this date the ripening process starts, thus it seems that the roughness decreases because the crops tend to bend a bit and the ears expend as well.

Soil roughness The soil roughness was only measured during early vegetation stages (until booting), when it still had an impact on the microwave signal. The actual method is the same as for the retrieval of the vegetation roughness. The board was only $50 \times 20cm$ (height $\times$ width) with a $2.5 \times 2.5cm$ raster. During the digitalization the point distance was only 1cm instead of 2cm.

Figure 3.10: Digitalisation of the vegetation profiles, using a point distance of 2cm. Exemplarily shown for 04/12/2006 (tillering) and 06/07/2006 (heading).

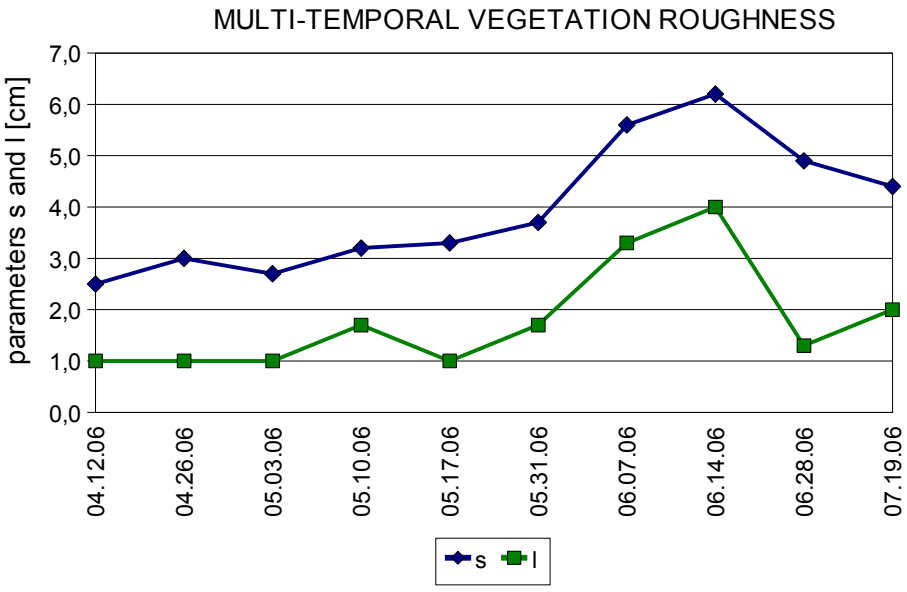

Figure 3.11: Standard deviation of vegetation height (s) and the correlation length (l) for wheat field 1 (2006).

Soil moisture content The soil moisture content was measured within the ground truth campaign. Thus, it was measured on the basis of weekly intervals and not in overflight synchrony with ERS-2 or Envisat ASAR, which was hard to realize because acquisitions (Envisat ASAR) were never really confirmed and the overflight times in the ascending orbits were rather late at night. Gravimetric (GSM) and volumetric (VSM) soil moisture content were measured at 0-5cm depth using a cylinder with a defined volume of $169cm^3$. Soil samples were dried at 110C for 24 hours in an oven.

$$GSM(\%) = \frac{W_W - W_D}{W_D} \times 100 \tag{3.7}$$

and

$$VSM(\%) = GSM(\%) \times DBD(gcm^{-3}) \tag{3.8}$$

With W_W the weight of wet soil, W_D the weight of dry soil and the dry bulk density (DBD) defined as the weight of the dry soil per unit volume.
In order to obtain information about the soil moisture content during the overflight times regression analysis between the measured soil moisture and the actual precipitation, which was measured hourly by a climate station were calculated. Best regression results ($R^2 = 0.7$) were found when taking a five days precipitation sum before the actual overflight time. Because soils are rather homogeneous in Klein-Altendorf this relationship was used for all experimental fields.

Agricultural management All experimental fields were normally managed by the local farmer, who reported the treatments within the *AGRO-NET*. Pest management was done around nine times within the vegetation period, starting from end of March. Fertilization was applied around three times from March until June.

3.1.3 Additional information sources

Winter wheat cultivation

Within two extensive field campaigns in the years 2005 and 2006 the agricultural use of the region was mapped, whereby the main interest were winter wheat fields. Fig. 3.1 shows the mapped area. In 2005 665 and in 2006 435 winter wheat fields were mapped within the research region.

Climatic data

During the vegetation growth periods 2005 and 2006 (April until August) a mobile climate station from the company ThiesClima[1] was placed within one of the research fields. The station measured in hourly intervals the precipitation [l/m^2], air temperature [°C], soil temperature [°C], relative humidity [%] and leaf wettness [%]. The measured variable precipitation was summed up for the whole day [$l/m^2/d$], while all others were averaged and respectively the

[1] www.thiesclima.com

minimum and maximum were calculated for that day. During the rest of the years (autumn, winter) a stationary climate station, belonging to the research farm measured all above mentioned parameters and the global radiation [W/m^2], as well as the daily sunshine hours. When comparing the two different climate stations the agreement between both was very good. Fig. 3.12 displays the measured variables precipitation -as monthly sum [mm]- and air temperature -averaged daily temperature [°C]- for the years 2005, 2006 and a long year average, from 1956 to 2000. When comparing these two years with each other and the long year average it becomes obvious that for both years temperatures from April until August were higher than the average. In 2006 the precipitation was higher as the average until July, where it decreased to around half of the long time average. For 2005 precipitation was rather high in April and May, and in the months June and July slightly lower than average. In general the year 2005 was wetter between the months March until July (2005: $\sum 314.6mm$, 2006: $\sum 278.0mm$ and average: $\sum 273.9mm$) then the average and 2006. The temperatures of both years were nearly the same, thus being in general higher than the average (2005: $\phi 13.4°C$, 2006:$\phi 13.2°C$ and average $11.7°C$).

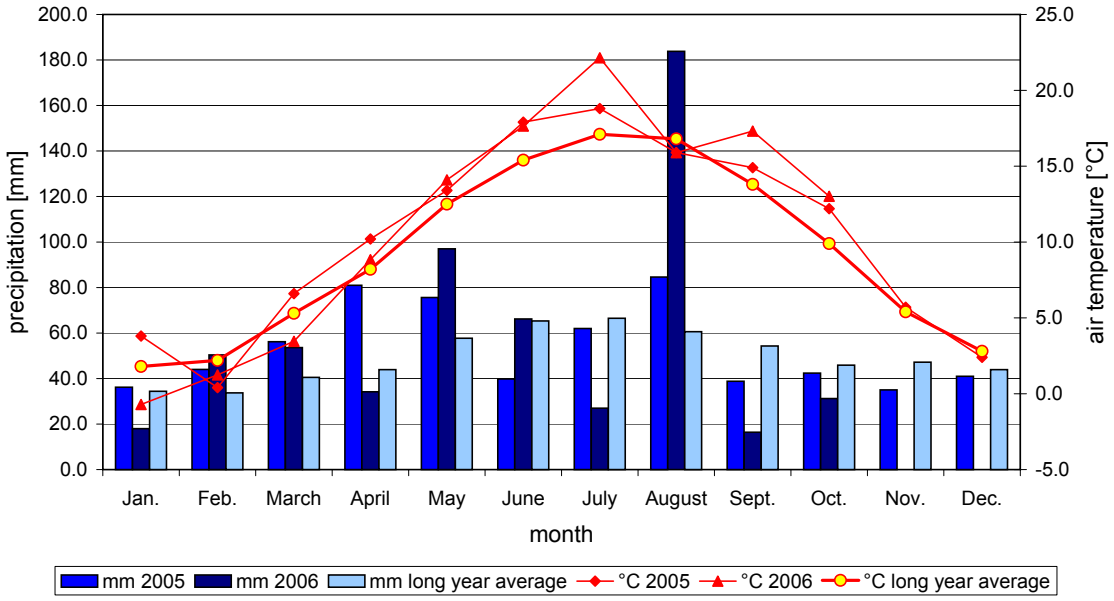

Figure 3.12: Monthly precipitation [mm] and averaged daily temperature [°C] values in 2005 and 2006 in comparison to a long year average

3.2 Remote sensing data and preprocessing

Within the study different kinds of multispectral sensors and SAR systems were used. Multispectral and SAR data need a slightly different kind of preprocessing, whereby especially the final step of geometric correction to one single reference image, is important when using different multi-temporal and multi-sensoral data.

3.2.1 Multispectral data

For generating a dense time-series different multispectral sensors were used, however still the amount of cloudy images and therefore not usable ones was rather high. In 2005 from seven planned Landsat 5 TM and three QuickBird imagery only two Landsat 5 TM and two QuickBird imageries were actually usable. In the year 2006 16 SPOT data sets were planned, whereby 8 imagery could only be taken into consideration. This illustrates the problem of cloud coverage in the mid-latitudes under current repetition rates. Tab. 3.4 lists all used data for both research years.

Table 3.4: Multispectral data used in 2005 and 2006 with the according phenological stage as reference.

multispectral sensor	date	phenological stage		multispectral sensor	date	phenological stage
ASTER	04/03/2005	elongation		SPOT 4	05/04/2006	elongation
QuickBird	04/22/2005	elongation		SPOT 2	05/10/2006	elongation
Landsat 5 TM	05/05/2005	elongation		SPOT 5	05/11/2006	elongation
SPOT 5	05/12/2005	elongation		SPOT 5	06/24/2006	flowering
Landsat 5 TM	05/28/2005	heading		ASTER	07/03/2006	ripening
QuickBird	06/20/2005	flowering		SPOT 5	07/14/2006	ripening

ASTER The *A*dvanced *S*paceborne *T*hermal *E*mission and *R*eflection Radiometer (ASTER) was launched in December 1999 by the NASA (Earth Observing System) and Japans Ministry of Economy Trade and Industry (METI) on the satellite Terra[2]. The ASTER instrument consists of the three separate instrument subsystems: Visible and Near Infrared (VNIR), Shortwave Infrared (SWIR) and Thermal Infrared (TIR) (Tab. 3.5).

Landsat 5 TM The Landsat *T*hematic *M*apper 5 (TM) was launched 1984 in a series of satellites by the NASA. A commercialization of the Landsat system was realized from 1985 until 2001 (Space Imaging EOSAT later, Space imaging), afterwards control was returned to the federal government again. The TM sensor primarily detects reflected radiation from the Earths surface in the visible, through the mid-IR, into the thermal-IR portion of the electromagnetic spectrum. Sixteen detectors for the visible and mid-IR wavelength and four detectors in the thermal-IR band provide information on each active scan[3] (Tab. 3.5).

[2]http://asterweb.jpl.nasa.gov
[3]http://edc.usgs.gov/guides/landsat_tm.html

QuickBird The QuickBird satellite was launched in 2001 by the company DitialGlobe from California, USA. It is designed to efficiently and accurately image large areas with a capability of acquiring over 75 million square kilometres of imagery data annually[4] (Tab. 3.5).

SPOT SPOT Image operates the used SPOT 2, SPOT 4 and SPOT 5 satellites and is the worldwide distributor of geographic information products and services derived from the SPOT Earth observation satellites. Data were made freely available within the *O*ptimizing *A*ccess to *S*POT *I*nfrastructure for *S*cience (OASIS), a European project coordinated by CNES and financed by the European Commission[5] (Tab. 3.5).

3.2.2 Preprocessing of the multispectral data

All data were ordered in the basic 1B processing level with radiometric and geometric corrections already applied. The radiometric correction included the application of a linear model for compensating instrument effects and normalizing CCD detector response. The geometric correction applied, accounts for systematic effects, including: panoramic distortion, Earths rotation and curvature, and variations in the satellites orbital altitude relative to the reference ellipsoid[6]. For further processing and analysis the ASTER SWIR (4-9) and TIR bands (10-14), as well as Landsat 5 TM TIR band (6) were not considered any more, either because of their geometric resolution or their spectral characteristics.
Before actual image analysis all data have to undergo a standardized preprocessing (radiometric and geometric correction), because they still feature systematic defects or undesirable sensor characteristics (Lillesand & Kiefer, 2000; Richards, 1993; Schowengerdt, 1997). Especially for multi-temporal or multi-sensoral image analysis radiometric corrections, considering sensor and atmospheric influences, and geometric corrections (the allocation to an uniform coordination system) are very important processing steps. A schematic overview of the different preprocessing steps is given in Fig.3.14.

Radiometric correction

In general the radiometric correction can be divided into two different steps. The first is the conversion of the sensors digital numbers (DN) to at-sensor radiance, whereby sensor calibration information is needed (*sensor calibration*). In a second step the at-sensor radiance is transformed into the radiance at the earth surface, considering the atmospheric condition at the time and location of image acquisition (*atmospheric correction*).

Sensor calibration
In this step the DN values are converted into radiances by the use of calibration coefficients (cal_gain_b and cal_offset_b) with units of radiance-per-DN, which are normally given by the

[4] http://www.digitalglobe.com
[5] http://www.spotimage.fr
[6] http://spot4.cnes.fr/spot4_gb/acquisit.htm

Table 3.5: Technical data of the used multispectral sensors

	Aster	Landsat 5 TM	QuickBird	SPOT 2 & SPOT 4	SPOT 5
inclination [°]	98.3	98.2	98.0	98.7	98.7
orbit type	sun-syn.	sun-syn.	sun-syn.	sun-syn.	sun-syn.
orbit period [min]	98.88	98.9	98.0	101.4	101.4
orbit repetition [days]	16	16	3.5	26	26
incidence angles [°]	0	0	0-30	± 31.06	± 31.06
geometric resolution [m]	VNIR: 15 SWIR: 30 TIR: 90	VNIR: 30 TIR: 120	PAN: 0.6-0.73 VNIR: 2.4-2.9	PAN: 10 VNIR: 20	PAN: 5 VNIR: 10
spectral resolution [μm]	 0.52-0.6 (Green) 0.63-0.69 (Red) 0.76-0.86 (NIR) 1.6-1.7 (SWIR) 2.145-2.185 (SWIR) 2185-2.225 (SWIR) 2.235-2.285(SWIR) 2.295-2.365 (SWIR) 2.36-2.43 (SWIR) 8.125-8.475 (TIR) 8.475-8.825 (TIR) 8.925-9.275 (TIR) 10.25-10.95 (TIR) 10.95-11.65 (TIR)	0.45-0.52 (Blue) 0.52-0.6 (Green) 0.63-0.69 (Red) 0.76-0.9 (NIR) 1.55-1.75 (SWIR) 10.4-12.5 (TIR) 2.08-2.35 (SWIR)	0.45-0.52 (Blue) 0.52-0.6 (Green) 0.63-0.69 (Red) 0.76-0.9 (NIR)	 0.5-0.59 (Green) 0.61-0.68(Red) 0.78-0.89 (NIR) 1.58-1.75 (SWIR)	 0.5-0.59 (Green) 0.61-0.68 (Red) 0.78-0.89 (NIR)
image size [km^2]	60×60	185×185	64×64	60×60	60×60

sun-syn. = sun-synchronous

data provider. The coefficients can be applied to the pixel values in each band (DN_b) by:

$$\textit{at-sensor } L_b^s = cal\_gain_b \times DN_b + cal\_offset_b \tag{3.9}$$

producing band-integrated radiance values (L_b^s).

Atmospheric correction
Aim of the atmospheric correction is to remove wavelength dependent effects of the atmosphere, such as scattering and absorption processes and thus the various paths and the different components of sky radiance and path radiance (Lillesand & Kiefer, 2000; Mather, 1999).
The scattering processes occur when particles or gas molecules are present in the atmosphere and interact with and cause the electromagnetic radiation to be redirected from its original path. The magnitude of the scattering process depends on factors like the wavelength of the radiation, the abundance of the particles or gases, and the distance the radiation travels through the atmosphere. Three different scattering types are taking place:

- Rayleigh scattering, when particles are very small compared to the wavelength of the radiation
- Mie scattering, occurring when particles (e.g., dust, pollen, smoke, water vapour) are just about the same size as the wavelength of the radiation
- non-selective scattering, if particles (e.g., water droplets) are much larger than the wavelength of the radiation

Absorption, being the other main mechanism, causes molecules in the atmosphere to absorb energy at various wavelengths. For example absorbs Ozone, the "harmful" ultraviolet radiation, carbon dioxide strongly absorbs the far infrared portion of the spectrum and water vapour much of the incoming long wave infrared and shortwave microwave radiation. Through this, different radiance paths in the sun-earth-sensor system are possible, like the direct path (sun-earth target-sensor), or the path sun-atmospheric scattering-sensor, or paths like sun-foreign target-sensor (Mather, 1999).
Actual radiometric correction was done with ATCOR 2, a model developed by Richter (1990) for the atmospheric correction of flat terrain, assuming Lambertian reflection. Whereby, lookup-tables are used and not an exact physical modelling of the atmosphere. Main variable parameters within the model are:

- water vapour (between 0.8-4.75 g/cm^2)
- aerosol type (rural, urban, maritime, oceanic or desert)
- visibility (between 5-120km)
- height above sea level
- solar zenith angle
- incidence and azimuth angle for tilting sensors

- interpolated atmospheric functions

For each single image in a first step the fixed parameters: sensor (e.g., Landsat 5 TM, SPOT 5), acquisition date and time, solar zenith angle, average relief hight and incidence angle have to be configured. In a second step the sensor specific calibration factors were defined, either coming from the data header or from the image provider. In a third step the visibility during acquisition was specified, whereby the true visibility was chosen taken from *www.wetter-online.de*. In the last step the atmospheric condition and the aerosol type had to be defined. For Germany generally the aerosol type *urban* is recommended, due to the rather high site density in comparison to North America. Only the parameter atmospheric condition had to be varied until the modelled reflectance curves corresponded the reference curves from different land surface targets, e.g., water, urban, forest, grass. For the actual correction the option *no haze removal* was taken, because the haze removal seemed to introduce artefacts into the images. In a last step a bi-directional reflectance factor (BRDF) correction was performed.

Geometric correction

Although data were ordered in level 1B format, whereby systematic geometric corrections have already been applied, images still have a location bias due to topographic effects and are not all in the same coordination system. Therefore, an image-to-image geometric correction method was used, projecting all images (as well as later the SAR data) to one reference image into the coordinate system UTM WGS 84, zone 32 north. Images were ortho-rectified using a digital elevation model (30m resolution) and a 3rd order polynomial transformation with rmse < 0.5. The chosen resampling strategy was cubic convolution, whereby the new pixel values are computed from weighting 16 (4×4) surrounding pixels, which smoothest the image but on the other hand slightly modifies the pixel values.

3.2.3 SAR data

Different kind of SAR sensors were used within the study for analysing the potential of different polarizations and incidence angles for applications with an agricultural background. Due to the SAR advantage of being weather (cloud) independent the time-series for ERS-2 are actually rather dense. For Envisat ASAR data ordering problems occurred, especially for some tracks and *A*lternating *P*olarization (AP) modes, due to conflicts with commercial and other scientific users, thus not all ordered images were actually acquired. All data used within the study are listed in Tab. 3.6 (ERS-2 SAR) and 3.7 (Envisat ASAR), together with the phenological stages and other data relevant features.

ERS-2 The second *E*uropean *R*emote *S*ensing satellite (ERS-2) was launched by the *E*uropean *S*pace *A*gency (ESA)[7] in 1995. The used *Active Microwave Instrument* (AMI) can operate in three modes, whereby especially the *Synthetic Aperture Radar* (SAR) is of interest for land applications (Tab. 3.8).

[7] http://www.esa.int/

Table 3.6: ERS-2 SAR data used in 2005 and 2006 with the according phenological stage.

2005					2006			
date	phenological stage	orbit	track		date	phenological stage	orbit	track
04/05/2005	elongation	desc.	108		04/06/2006	elongation	desc.	337
04/15/2005	elongation	asc.	258		04/25/2006	elongation	desc.	108
04/21/2005	elongation	desc.	337		05/11/2006	elongation	desc.	337
05/10/2005	elongation	desc.	108		05/30/2006	heading	desc.	108
05/26/2005	heading	desc.	337		06/09/2006	heading	asc.	258
06/05/2005	heading	asc.	487		07/20/2006	ripening	desc.	337
06/24/2005	flowering	asc.	258					
07/19/2005	ripening	desc.	108					
08/04/2005	ripening	desc.	337					

asc.= ascending; desc.=descending

Envisat The *Envi*ronmental *sat*ellite (Envisat)[8] is an Earth-observing satellite launched (2002) by the ESA. It carries nine instruments, whereby several of them are advanced versions of the six instruments from the earlier ERS-1 and ERS-2 missions.
The *Advanced Synthetic Aperture Radar* ASAR instrument used, has an active phased array antenna, allowing the independent control of phase and amplitude of the transmitted radiators from different regions of the antenna surface and a weighting of these received signals. This allows the operation in different modes: the conventional stripmap SAR or the ScanSAR mode (used). The stripmap SAR allows the Image Mode operating in seven predetermined swaths with either vertically or horizontally polarized radiation. The Wave Mode uses the same swaths and polarizations as the Image Mode, but only images small areas at regular intervals (5×5km vignettes). For the ScanSAR Modes the normal limitation of the stripmap SAR (narrow swath) is overcome by widening the use of an antenna beam, which is electronically steerable in elevation. Operational modes are the Wide Swath Mode and the Global Monitoring Mode, as well as the used *A*lternating *P*olarization (AP) mode. The distinctive feature of the AP mode is that instead of scanning between different elevation sub-swath, scanning is performed between two polarizations (HH, VV) (Fig. 3.13). Additionally cross-polar modes are possible, where the transit pulses are either H or V, with an alternatively receiving in H or V (see Tab. 3.8), thus allowing the recording of the cross-polarizations HV or VH.

3.2.4 Preprocessing of the SAR time-series

SAR data were ordered in the *PR*ecision *I*mage (PRI) format, which are multi-looking (speckle-reduced) ground range images. These images are calibrated and corrected for the SAR antenna

[8] http://envisat.esa.int/

Table 3.7: AP Envisat ASAR data (HH and HV) used in 2005 and 2006 with the according phenological stage.

date	phenological stage	orbit	track	swath	incidence angle [°]
04/12/2005	elongation	desc.	208	6	39.1-42.8
07/10/2005	ripening	asc.	487	2	19.2-26.7
07/13/2005	ripening	desc.	029	3	26.0-31.4
07/22/2005	senescence	asc.	158	3	26.0-31.4
04/16/2006	elongation	asc.	487	2	19.2-26.7
05/02/2006	elongation	desc.	208	6	39.1-42.8
05/21/2006	elongation	asc.	487	2	19.2-26.7
06/06/2006	fruit development	desc.	208	6	39.1-42.8
06/25/2006	fruit development	asc.	487	2	19.2-26.7

asc.= ascending; desc.=descending

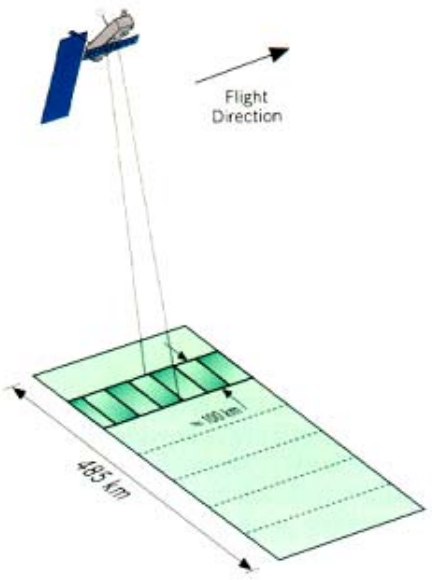

Figure 3.13: Alternating polarization mode

http://envisat.esa.int/handbooks/asar/CNTR.htm

Table 3.8: Technical data of the used SAR systems.

	ERS-2 SAR	Envisat ASAR
inclination [°]	98.5	98.5
orbit type	sun-syn.	sun-syn.
orbit period [min]	100	101
orbit repetition [days]	35	35
frequency [GHz]	3.5	5.331
wavelength [cm]	5.66	5.62
polarization	VV	VV/VH
		HH/HV
		VV/HH
incidence angles [°]	19.2-26.7	15.0-22.9
		19.2-26.7
		26.0-31.4
		31.0-36.3
		35.8-39.4
		39.1-42.8
		42.5-45.2
pixel spacing [m]	12.5	12.5
swath width	100	56-105*

*= incidence angle dependent; sun-syn. = sun-synchronous

pattern and range-spread loss, thus SAR backscatter can be derived from the product for geophysical modelling, but no correction for terrain-induced radiometric effects is applied. The images are not geometrically corrected, and terrain distortion (foreshortening and layover) has not been removed (for further detail please refer to: Laur (1998); Oliver & Quegan (2004); Woodhouse (2006)). A detailed overview of all the preprocessing steps is given in Fig. 3.14.

Calibration

In general for the calibration of PRI images a flat terrain is considered, thus the incidence angle α only depends on the ellipsoid and varies from near to far range (e.g., ERS-2: from 19.5° to 26.5°). Since the study site is rather flat (inclination around 0.5°-1°) the assumption is valid and no further topographic normalization has been performed. The SAR backscattering coefficient σ^0 is related to the SAR brightness β^0 and the according incidence angle α. Whereby, the SAR brightness β^0 is defined as the property of an image in which the strength of the SAR reflectivity is expressed as being proportional to a digital number.

$$\sigma^0 = \beta^0 \times sin\alpha \tag{3.10}$$

Due to the assumption of a flat terrain PRI images are coded as SAR brightness β^0. The pixel intensity values are directly proportional to the brightness by:

$$[DN]^2 = constant \times \beta^0 = constant \times \frac{\sigma^0}{sin\alpha} = constant(\alpha) \times \sigma^0 \tag{3.11}$$

with $constant(\alpha)$ defined as:

$$constant(\alpha) = K \times \frac{sin\alpha_{ref}}{sin\alpha} \tag{3.12}$$

whereby, K is the calibration constant and α_{ref} the mid-range incidence angle, e.g., for ERS SAR 23 (for further details please refer to: Laur (1998)).

Co-registration

Co-registration between images of the same track, having the same image geometry, was performed with the software tool BEST[9]. In an initial step the initial registration is performed using the satellite orbit parameters. In a second step a coarse registration is carried out using a cross-correlation operation on a series of "cells", which are defined across the images. In the final step a further fine registration is carried out by maximizing the complex coherence between the images for a series of "cells". This allows a further improvement of the cross-correlation function. By doing this one actually gains a transfer function for the co-registration process for each point.

Speckle filtering

A distinctive characteristic of SAR images is the so called speckle or random "salt and pepper" noise, due to the actual distribution of scatterers at the level of wavelength. Even on a bare

[9] http://earth.esa.int/services/best/

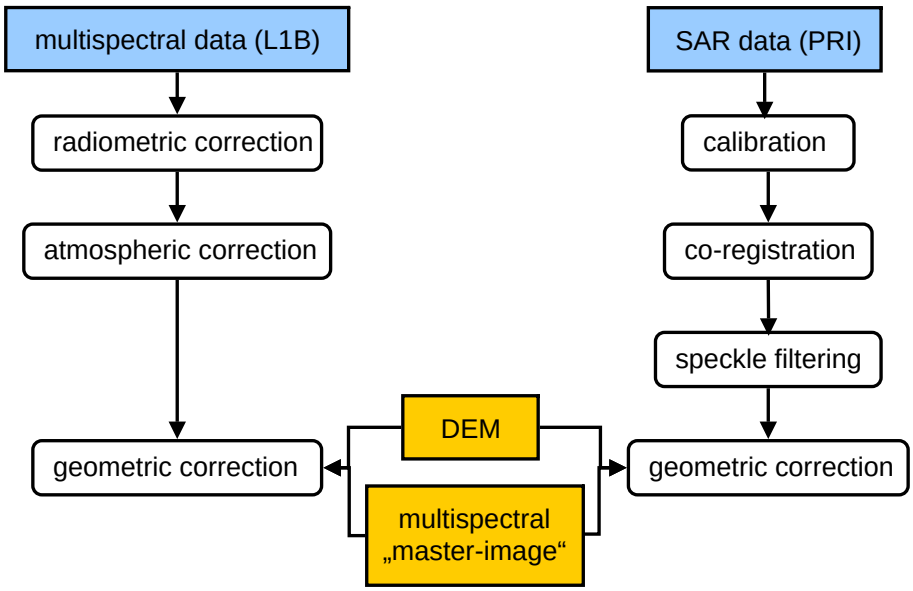

Figure 3.14: Preprocessing chain for the multispectral and SAR data

field, appearing homogeneous on the ground, characteristics like surface roughness, normally being statistically homogeneous over the field, cause small variations of the scatterer location (Woodhouse, 2006). By ordering PRI images the speckle effect was already reduced through multi-looking, whereby the single looks, corresponding to multiple measurements of the same resolution cell, are averaged. In order to further decrease the speckle the enhanced Frost Filter (5×5 window) was applied, trying to average across homogeneous areas but preserving boundaries. This filter assumes that speckle is a multiplicative noise and belongs to the category of adaptive-weighted-mean filters (Lopes et al., 1990).

Geometric correction

After co-registration of the same tracks, the geometric correction of the different acquired tracks to a single “master” image and the projection to the coordinate system (UTM, WGS-84, zone 32 north) is necessary. As a *master* image a multispectral Landsat 5 TM mosaic from the year 2000 was used (same as for the multispectral data). The correction was performed by using the orbit information, ground control points and a digital elevation model. For the resampling the cubic convolution method was chosen.

4 BIOPHYSICAL PARAMETER ESTIMATION

In the following chapter firstly the retrieval of biophysical parameters using multispectral data will be analysed. Then the use of SAR data for biophysical parameter estimation will be evaluated, followed by a combined approach using both multispectral and SAR data. In Fig. 4.1 a rough overview of the different analysed retrieval methods is given. From the multispectral data different vegetation indices are calculated, from the SAR data the phenological stage heading/flowering, the local incident angle and the backscatter coefficients are derived. This information is then used with additional information about precipitation for empirical and semi-empirical parameter modelling, either on the basis of one sensor system or by fusing the information from the different sensors. Previous to all analysis steps there are some additional considerations necessary for generating multispectral and SAR time-series. For the multispectral data only cloud free images were chosen, leaving out five Landsat 5 TM and five SPOT images within the years 2005 and 2006 (already not mentioned in Tab. 3.4). For the SAR images precipitation is an important feature, because intense rain can hamper the image interpretation. For each image the precipitation, measured at the local climate station, was considered. E.g., ERS-2 acquisitions on the 06/30/2005 and 06/15/2006 are not considered for the analysis and not mentioned in Tab. 3.6.
Another important feature is the precision of the measured ground truth data. Therefore the standard deviation of each field and the multi-temporal dynamics of each parameter were considered for each image (see section Ground Truth Data 3.1). Due to variabilities, e.g., the ERS-2 image from 05/26/2005 was only taken into account for field 1, from 05/30/2006 only for fields 2, 3 and 4 and from 06.15.2006 only fields 1 and 2. Concerning the multispectral data for the Landsat 5 TM image (05/28/2005) field 2 was not considered and during the year 2006 field 2 for the SPOT 5 image (06/24/2006) and field 4 for the ASTER image (07/03/2006).

4.1 Biophysical parameter extraction using multispectral data

Within this section the main question analysed was:

Is it possible to derive the values of important crop variables from various remote sensing data?

Whereby, in more detail the following challenges were of concern:

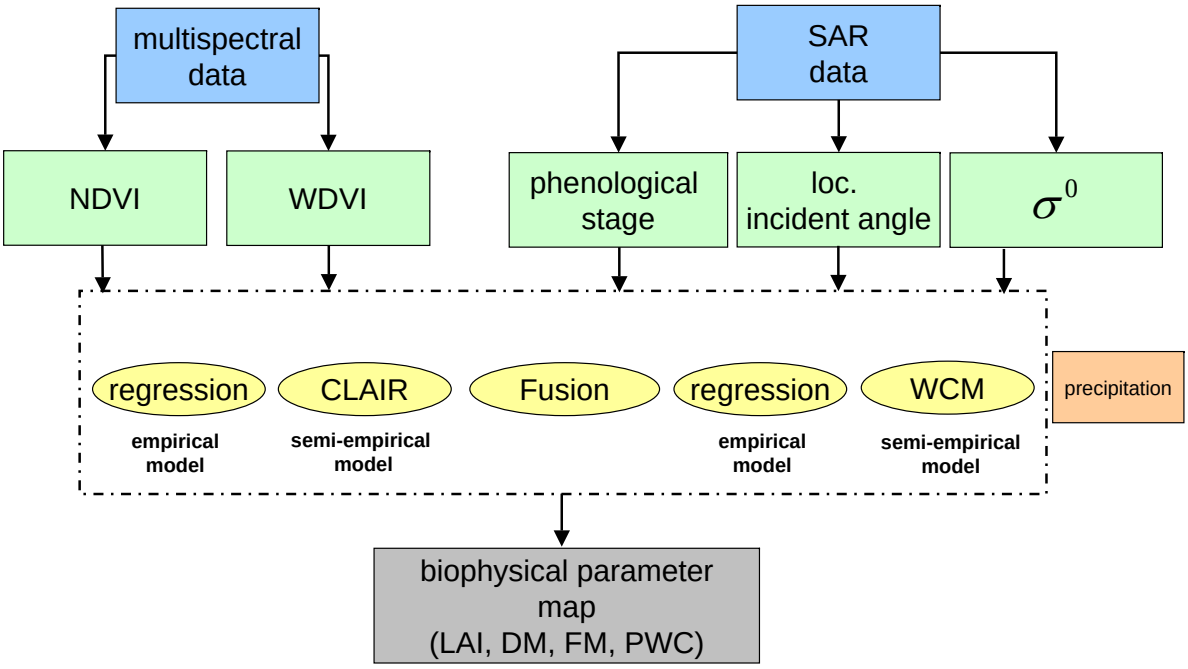

Figure 4.1: Overview of the biophysical parameter retrieval.

- to intercalibrate vegetation indices of different multispectral sensors in order to correct the biases, introduced by differences in the relative spectral response functions (section 4.1.1)

- to evaluate the performance of the CLAIR model for the retrieval of different biophysical parameters (section 4.1.3)

- to invert the used models for regional biophysical parameter extraction (section 4.1.3)

As a first analysis step, different vegetation indices (VI) were calculated. The most important ones for further investigations were the *N*ormalized *D*ifference *V*egetation *I*ndex (NDVI) (Rouse et al., 1978) and the *W*eighted *D*ifference *V*egetation *I*ndex (WDVI) (Clevers, 1988, 1989; Clevers et al., 2002). As a last processing step, average VIs were calculated for all analysed fields and the parcel boundaries were buffered with 30m zones. Buffering was necessary in order to avoid image pixels (mixed pixels) near parcel boundaries, partly covering other classes, e.g., roads or other agricultural fields. Finally the ground truth information and satellite data were matched into pairs by considering the acquisition dates trying to minimize the time gaps.

4.1.1 Normalized Difference Vegetation Index

For the retrieval of the different biophysical parameters using the NDVI the following linear model was chosen:

$$E(Y/X = x) = \alpha + \beta x + \epsilon \tag{4.1}$$

with Y the biophysical parameter and X the NDVI, the unknown coefficients α and β and ϵ the error term. Tab. 4.1 shows the adjusted coefficients and the coefficient of determination (R^2) for the independent validation process. For modelling the leave-one-out cross-validation method was chosen, considering all available datasets within the study years 2005 and 2006. In a second step it was tried to minimise the differences between the VIs, introduced by variations in the RSR functions of the respective used remote sensing sensors, as differences in the VI have an impact on the retrieval performances.

Table 4.1: Fitted coefficients and coefficients of determination (validation) for the linear empirical biophysical parameter modelling using the NDVI.

	FM	DM	LAI	PWC
α	2.79	0.83	-1.76	0.47
β	1.05	0.06	3.82	2.82
R^2	0.67	0.42	0.46	0.70

(FM= fresh biomass [kg/m^2]; DM= dry biomass [kg/m^2]; LAI= leaf area index [m^2/m^2]; PWC= plant water content [kg/m^2])

Cross-calibration of the NDVI

Several studies already have analysed the discrepancies in VIs due to different relative spectral reflectance curves of the remote sensing sensors and attempted to correct them by using empirical cross-calibration methods. Based on a multispectral sensor simulation using hyperspectral HyMap data these resulting discrepancies were analysed and are shown here for the NDVI. For the sensor simulation only the spectral characteristics of the sensors were simulated by constant side-parameters like, e.g., geometric features (Franke et al., 2006). As a prerequisite for data simulation, the RSR values of each multispectral band were fitted to the according wavelengths of the 126 hyperspectral HyMap bands. Therefore, each HyMap center wavelength was linked with the mean RSR value (in the range of FWHM of the hyperspectral band) of the simulating band. A comparison between the generated RSR functions, which were calculated by the mean RSR value, maximum RSR value and the actual RSR value at HyMap's center wavelength, showed that the mean RSR value approach best corresponds to the true RSR functions of the sensors. For the simulation process the 126 HyMap reflectance values of each pixel were multiplied by the 126 wavelength corresponding RSR values of the simulating band. The sum of these products is divided by the sum of the 126 band-specific RSR values. For a multispectral sensor simulation, each band must be simulated according to the following equation, where R_{sim_b} is the simulated pixel reflectance value of the simulated band, R_i is the pixel reflectance value of the HyMap band and $rsr_{b,i}$ is the RSR value of the simulating band at each HyMap corresponding wavelength:

$$R_{sim_b} = \frac{\sum_{i=1}^{126} R_i \times rsr_{b,i}}{\sum_{i=1}^{126} rsr_{b,i}} \tag{4.2}$$

The results were simulated multispectral bands, providing spectral information similar to those of the original sensor bands. Nevertheless, differences to original data caused by different

spatial resolution or sun/sensor/target geometry still exist. This fact has to be considered by the validation of the simulation (Franke et al., 2006).

In order to assess the accuracy of the simulation method an extensive evaluation was performed by the use of a real Landsat 5 TM image acquired on the same day, round about 1.30 hours after the HyMap acquisition. As a first indicator for the simulation quality, the differences in the overall statistics of the imagery were assessed regarding each band and the NDVI (Franke et al., 2006) (Tab. 4.2). The differences of both data sets do not vary significantly with values between 0.08% and 1.45% reflectance, by a mean of 0.79%. Highest congruence between real and simulated data was achieved in band 3, widest differences existed in band 7. Obvious was a general underestimation of pixel values in the simulated data. Validation of the NDVI differences between original and simulated Landsat 5 TM showed a very high congruence as well. Only marginal differences of 0.004 (0.62%) were found (Tab. 4.2). In general the validation by the comparison between real and simulated data demonstrated a good performance of the simulation method.

Table 4.2: Bandwise mean and standard deviation of reflectance (%), as well as NDVI for the real and simulated Landsat 5 TM scene and the reflectance plus NDVI differences between them.

	simulated Landsat		real Landsat		
	Mean	STDEV	Mean	STDEV	abs. difference
Band 1	4.0%	4.0%	5.0%	3.3%	0.96%
Band 2	6.9%	5.4%	7.7%	4.8%	0.80%
Band 3	7.5%	6.6%	7.5%	6.1%	0.08%
Band 4	34.3%	12.4%	34.8%	10.2%	0.46%
Band 5	18.0%	8.3%	19.0%	7.3%	1.01%
Band 7	11.0%	8.7%	12.4%	8.1%	1.45%
NDVI	0.635	0.277	0.639	0.258	0.004

The slight statistical differences had various reasons. For example, in comparison to the real Landsat 5 TM data set, the simulation did not consider variations due to different spatial resolution or sun/sensor/target geometry of the data. Additionally, different atmospheric corrections of real and simulated data, caused slight variations as well. Generally, variations in reflectance values occur due to different overpass time and off-nadir viewing. The last point is particularly critical for airborne sensors (Schiefer et al., 2006). In our study BRDF effects caused by different illumination and observation angles within the HyMap scene were negligible due to a south to north flight direction at noon. For the validation only areas in the centre of the HyMap scene (nadir) were taken into account. The solar illumination geometry was similar at Landsat 5 TM and HyMap overpass times (within 1.6). In the study of Kerekes & Landgrebe (1989), a simulation of a Landsat TM imagery on the basis of AVIRIS data was accomplished. The comparison showed an offset of the mean DN values from 0.4% up to 5.3%, according to the band. Teilet et al. (2001) compared Landsat 5 TM and 7 ETM+ and showed a spectral band difference in the range of 2-7% depending on the band leading to NDVI differences of 1% to 4% (mean 2.5%).

For a detailed trend analysis between the sensors the land cover types urban, forest, crops,

fallow and root crops were examined. In order to illustrate the differences between the two sensors (high minus low NDVI) in a comprehensible way, the sensors were all compared with each other (Fig. 4.2) taking the sensor with the higher NDVI values as a reference (x-axis) and the curves for a second order polynomial (R^2) as a reference and an indicator for the trend analysis (Heinzel et al., 2006).

The trend between SPOT 5 vs. Aster (Fig. 4.2 A) is relatively linear (second order polynomial, R^2=0.84), with a high difference and scattering increase towards the high NDVI range. The differences between SPOT 5 vs. QuickBird (Fig. 4.2 B) are the widest of all compared sensors. The trend is nearly linear (R^2=0.97), with a peak in differences towards the high NDVI values. The observed trend (Landsat 5 TM vs. SPOT 5) (Fig. 4.2 C) shows a concave shape (R^2=0.97), with a high dynamic in the NDVI range from 0.8 to 1 and a plateau in the differences trend in the middle NDVI range (0.2-0.6). The offset between Landsat 5 TM and Aster (Fig. 4.2 D) is described by a convex shape (R^2=0.76). It begins with relatively low differences (0.01) for low NDVI values (0.2), increasing up to 0.03 in the middle NDVI range (0.4-0.6) and finally decreasing again (0.01) for high NDVI values (0.8-1). The trend Landsat 5 TM vs. QuickBird (Fig. 4.2 E) occurs to have a convex form (R^2=0.86), when looking at the general overall trend, with the highest differences (0.03) in the middle NDVI range (0.4-0.6). The offset between QuickBird vs. Aster (Fig. 4.2 F) seems to be nearly even (R^2=0.73), with trend change in the differences.

In Tab. 4.3 the overall differences between the simulated sensors are shown. In regard to them the intercalibration performance will be assessed. Here again it becomes obvious that the differences between SPOT 5 and QuickBird are the highest, while the sensors QuickBird and Aster feature relatively similar NDVI. In general cross-calibration methods can be categorized

Table 4.3: MIN, MAX and MEAN NDVI differences and percentage difference between the simulated sensors. Percentage differences were calculated considering the NDVI mean of the corresponding sensor and the mean NDVI difference of the respective sensor pair.

sensor pairs	Min NDVI differences	Max NDVI differences	Mean NDVI differences	Differences (%)
SPOT5-Aster	-0.116	0.104	0.012	1.818
SPOT5-Landsat 5TM	-0.157	0.126	0.009	1.364
SPOT5-QuickBird	-0.064	0.099	0.045	7.087
Landsat 5TM-Aster	-0.088	0.091	0.003	0.472
Landsat 5TM-QuickBird	-0.074	0.083	0.013	1.970
QuickBird-Aster	-0.058	0.040	-0.002	-0.322

into two groups: calibration by regression (Miura & Yoshioka, 2006; Steven et al., 2003; Trishchenko et al., 2002) and by weighted averaging (Gao, 2000; Gittelson & Kaufman, 1998; Miura & Yoshioka, 2006). After analysing the trends in differences between the sensors it can be clearly seen that the cross-calibration cannot be easily modelled by linear regression as

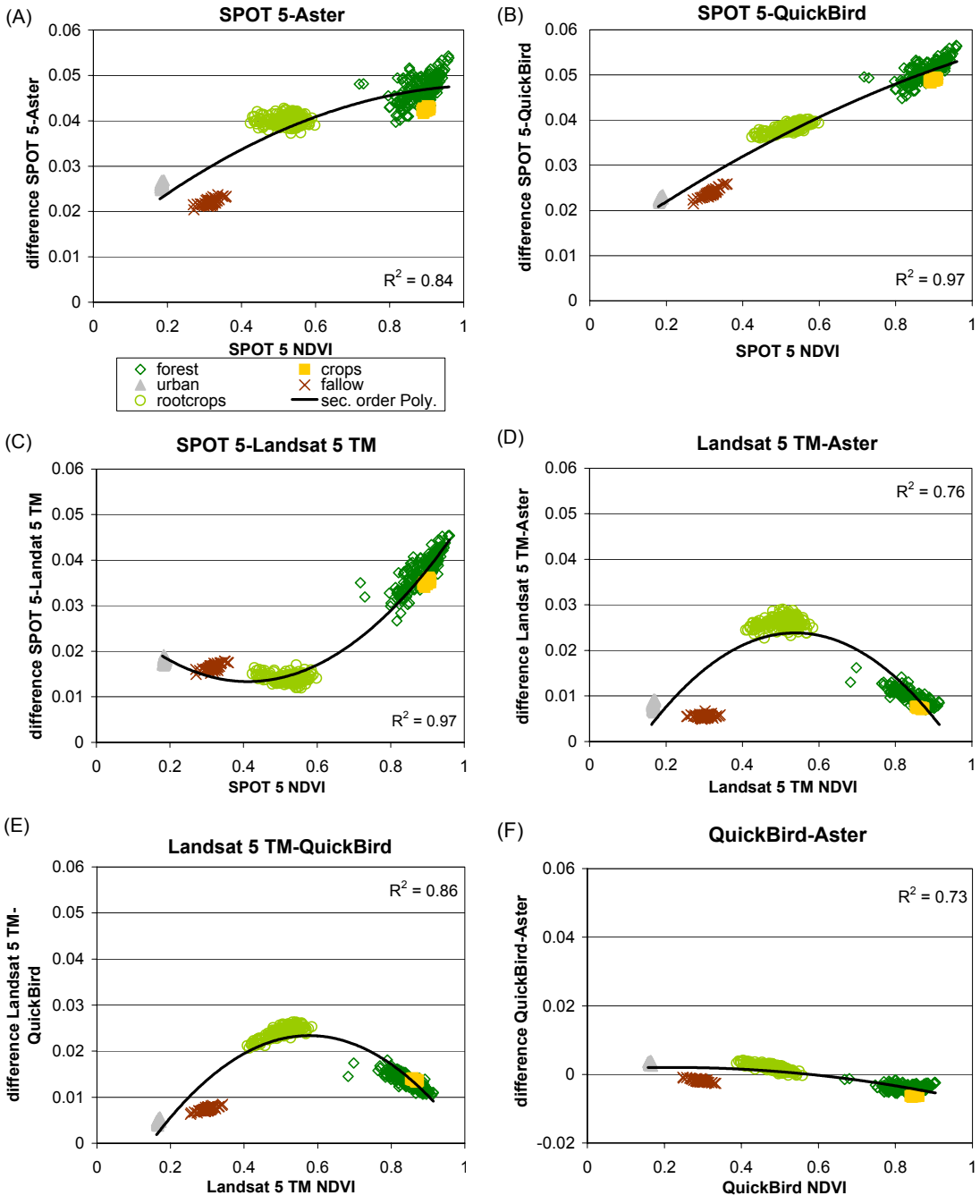

Figure 4.2: Absolute differences for high NDVI between the studied sensors.

Steven et al. (2003) did for only vegetated areas. In Miura & Yoshioka (2006) both methods (regression and weighted averaging) were tested and both performed well. But he stresses the fact, that when trying to cross-calibrate the whole NDVI range using the regression method, higher order polynomials are needed. This was also found in the study of Heinzel et al. (2007), when actually comparing the performance of second and sixth order modelling. When looking at the coefficients of determination, describing the relationship between two sensors for different polynomial orders, using a higher order, e.g., sixth order is also reasonable here (Tab. 4.4):

$$NDVI_{target} = a_0 + a_1 NDVI_{source} + ...a_{n-1} NDVI_{source}^{n-1} + a_n NDVI_{source}^{n} + \epsilon \quad (4.3)$$

where the subscripts *source* and *target* indicate the sensor to be translated as well as the reference sensor, *n* the order of the polynomial chosen, the different *a* the coefficients to be adjusted and ϵ the unexplained error term.
The translation coefficients for the sixth order regression translating one NDVI into the other are shown in Tab. A.1 in chapter **ATTACHMENT** (A). In order to asses the intercalibration performances of the sixth order modelling the overall minimum, maximum and mean NDVI differences and percentage differences between the simulated sensors were analysed (Tab. 4.5) and compared with the original values (Tab. 4.3).

When comparing the sensor differences they generally indicate a significant reduction after intercalibration. For example the highest improvement was found for SPOT 5 and QuickBird were the differences decreased from 7.087% to 0.152% or 0.161% (depending on the used target sensor). Overall, the best results with the smallest bias errors were obtained for translating Aster or QuickBird into SPOT 5. The magnitudes of error from translating the sensors into each other were between 0.152 and 0.606%. In general the approach delivers reasonable results, in comparison to the results from Steven et al. (2003) or Miura & Yoshioka (2006). In these studies a precision of 1-2% and 2% was achieved.

Table 4.4: Coefficients of determination for modelling the NDVI relationships using different polynomial orders.

sensor pair	polynomial order	R^2	sensor pair	polynomial order	R^2
SPOT 5 - ASTER	2	0.84	Landsat 5 TM - ASTER	2	0.76
	6	0.95		6	0.98
SPOT 5 - QuickBird	2	0.97	Landsat 5 TM - QuickBird	2	0.86
	6	0.99		6	0.99
SPOT 5 - Landsat 5 TM	2	0.97	QuickBird - ASTER	2	0.79
	6	0.98		6	0.93

Parameter retrieval using the intercalibrated NDVI

Due to the fact that the amount of data pairs (ground truth samples and multispectral images) is relatively small, a multi-sensoral biomass retrieval approach was realized after the intercal-

Table 4.5: MIN, MAX, MEAN NDVI differences and % difference between the original NDVI imagery and the cross-calibrated image for the sixth order polynomial modelling.

Differences were taken between the original sensor (org) and the cross-calibrated sensor (cross), which is still named after its origin.

org-cross	MIN	MAX	MEAN	%
SPOT 5-ASTER	-0.445	0.068	-0.001	-0.152
SPOT 5-LANDSAT 5 TM	-0.797	0.106	0.004	0.606
SPOT 5-QuickBird	-0.330	0.052	-0.001	-0.152
Landsat 5 TM-ASTER	-0.100	0.964	-0.002	-0.315
Landsat 5 TM-QuickBird	-0.084	0.910	-0.004	-0.630
Landsat 5 TM-SPOT 5	-0.107	0.954	-0.003	-0.472
QuickBird-ASTER	-0.057	0.031	0.001	0.161
QuickBird-Landsat 5 TM	-0.609	0.078	0.004	0.483
QuickBird-SPOT 5	-0.049	0.221	0.001	0.161
ASTER-Landsat 5 TM	-0.675	0.096	0.002	0.320
ASTER-QuickBird	-0.030	0.058	-0.001	-0.160
ASTER-SPOT 5	-0.067	0.225	-0.001	-0.160

ibration of the NDVI. Thus, as an intermediate step the sensors were all intercalibrated to the reference Landsat 5 TM. The actual empirical modelling of the biophysical parameters was performed by using the leave-one-out cross-validation method. Fig. 4.3 illustrates exemplarily the relationships between the modelled and observed fresh biomass [kg/m^2] using the intercalibrated NDVI and the original NDVI. The coefficients of determination (linear regression) for the different NDVI versions are shown in Tab. 4.6. It becomes clear that the NDVI intercalibrated to Landsat 5 TM using a sixth order polynomial model describes the relationship best. R^2 are the highest for all retrieved parameters and the scattering around the 1:1 line is the lowest.
Regarding the biophysical parameters the best performance was found for the fresh biomass and the plant water content, while the LAI retrieval performance was rather poor. In general all findings are still relatively poor compared to the literature. Especially, concerning the LAI modelling were Ustin (2004) achieved R^2 of around 0.9. However, in their study it became not clear if independent validations of the models were performed, which of course has an impact on the gained R^2.

Table 4.6: Fitted coefficients and coefficients of determination (validation) for the linear empirical biophysical parameter modelling using the intercalibrated NDVI.

	FM	DM	LAI	PWC
α	3.03	0.83	-1.61	0.79
β	0.72	0.06	3.71	2.50
R^2	0.73	0.54	0.53	0.72

(FM= fresh biomass [kg/m^2]; DM= dry biomass [kg/m^2]; LAI= leaf area index [m^2/m^2]; PWC= plant water content [kg/m^2])

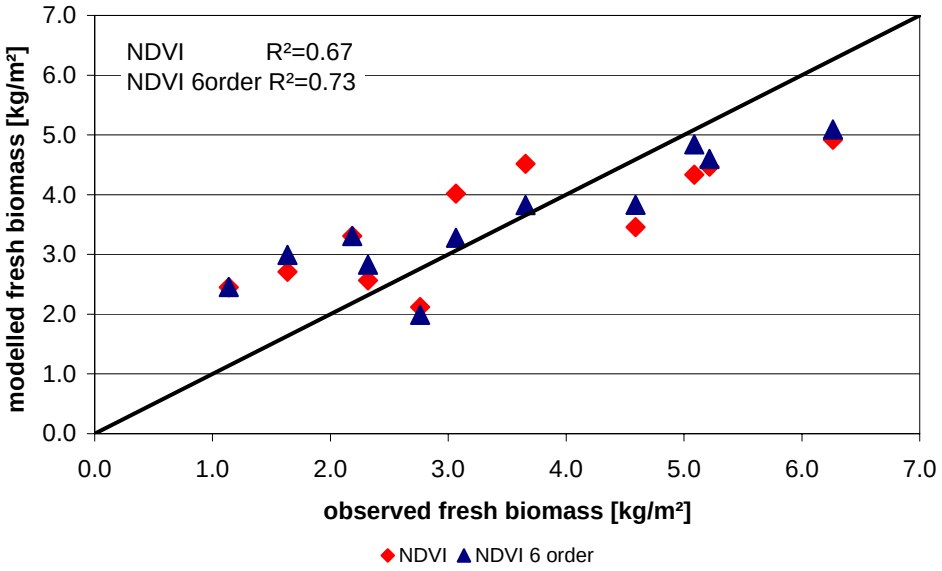

Figure 4.3: The relationships between observed versus modelled fresh biomass [kg/m^2], using the different NDVI modelling approaches.

4.1.2 Other indices and empirical regressions

On the basis of the ground truth measurements and the remotely sensed information also the relationships between different VIs and the biophysical parameters were analysed by considering different empirical modelling approaches, using the leave-one-out cross-validation method.
Indices analysed next to the NDVI, were the *P*erpendicular *V*egetation *I*ndex (PVI) (Jackson, 1983):

$$PVI = 0.647 \times NIR - 0.763 \times red - 0.02 \tag{4.4}$$

which uses a statistical relationship to remove effects of soil background on vegetation estimates. The *S*oil-*A*djusted *V*egetation *I*ndex (SAVI) (Huete, 1988) also accounts for the soil background using an adjustment factor ($L = 0.5$) for increasing the sensitivity, when separating canopy and soil at low ground cover amounts.

$$SAVI = \frac{NIR - red}{NIR + red + L}(1 + L) \tag{4.5}$$

And finally a simple ratio of NIR and red reflectance (NIR/red) was evaluated.
The relationships of these VIs and the biophysical parameters were analysed using different empirical regressions, e.g., linear, exponential, polynomial, power.
The Fig. 4.4 (A) illustrates exemplarily the power regression modelling using the different VIs for fresh biomass, whereby the PVI is not displayed, because retrieval results were out of range. The SAVI actually failed modelling the fresh biomass, the IR/RED VI had some problems in the higher biomass range, saturating there. The NDVI modelled the fresh biomass best with the lowest scattering for the whole biomass range. Concerning the dry biomass modelling using the SAVI outperformed the achievement, using the NDVI concerning the scattering and the R^2. This was also for the biophysical parameter LAI. For the plant water content modelling by the use of the NDVI was most explicit. Independently of the biophysical parameter and the empirical model used, the NDVI however was the most suitable for all cases.

Different empirical models using the NDVI are shown exemplarily for the modelled versus the observed fresh biomass in Fig 4.4 (B). Generally, it seems that for low biomass, smaller than 4 kg/m^2, there was an overestimation, while for higher biomass there was rather an underestimation. Overall, the power regression seems to model the fresh biomass over the whole range with the highest accuracy ($R^2 = 0.77$).

Regarding the dry biomass the exponential model in combination with the SAVI, for LAI the

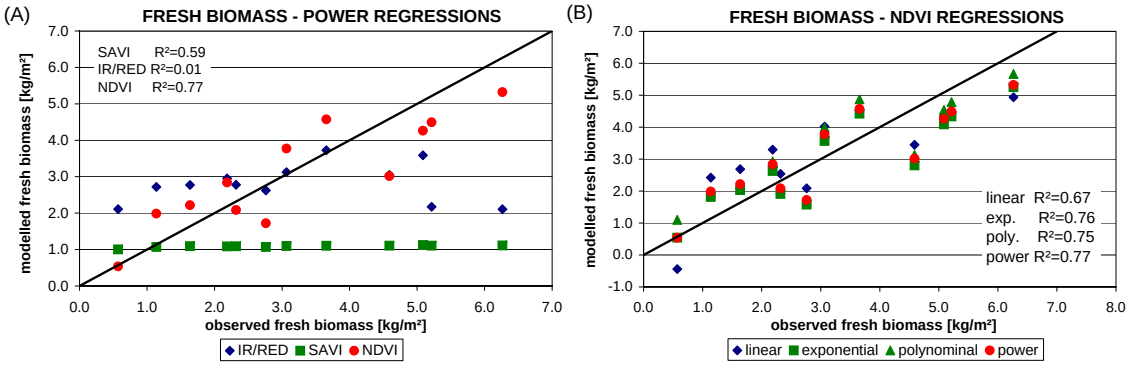

Figure 4.4: (A) Modelled versus observed fresh biomass for modelling using different VI. (B) Modelled versus observed fresh biomass for different regression models using the NDVI.

polynomial model using the SAVI and for the water content the NDVI in combination with the power modelling were most suitable. The best performance for all biophysical parameters was achieved using the NDVI and the power modelling approach. The achieved R^2 is within the lower range of the results in the reviewed literature, however, some of them propose a more linear relationship.

4.1.3 Semi-empirical CLAIR model

In an approach to overcome the negative effects of empirical modelling, meaning the problems when trying to transfer the findings in time or space and the relative poor modelling performance, also the modelling performance using a semi-empirical model (CLAIR) was assessed. The CLAIR model was actually designed for the parameter LAI. Thus, first of all the hypothesis that the parameters fresh and dry biomass as well as the plant water content are correlated with the LAI and therefore it might be possible to also retrieve them by the use of the CLAIR model, has to be analysed.

Fig. 4.5 illustrates the relationship between the WDVI and the LAI for the research years 2005 and 2006 during the whole growing season. The parameters α and $WDVI_{\infty}$ were fitted with the leave-one-out cross-validation method.

The results for all biophysical parameters analysed were not significant with coefficients of determination (R^2) below 0.5, only for the LAI R^2 was slightly above 0.5. Other research findings (Clevers et al., 2002) have also assessed this and suggested to divide the time-series into two different phenological phases, before heading/flowering and after heading/flowering. This was

done by either using the ground truth information or by analysing the ERS-2 time-series and extracting the heading/flowering dates (section 4.2.4). Fig. 4.5 displays the relationship between the WDVI and the LAI when dividing the time-series (blue boxes=before heading, red triangle=after heading). For all assessed biophysical parameters the extraction accuracy increased significantly, when dividing the time-series into different distinctive phenological stages. Retrieval results were sounder than for the previous modelling approaches, using different empirical models and VIs. However, due to differences in the RSR functions of the used multispectral sensors a intercalibration of the WDVI is performed in order to improve the modelling accuracy.

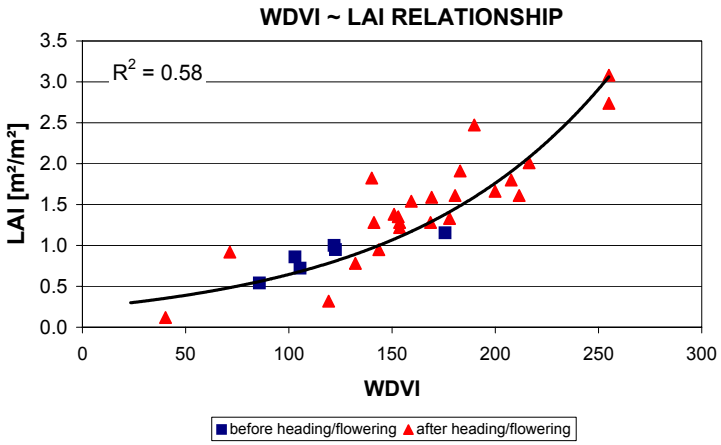

Figure 4.5: WDVI-LAI relationship.

WDVI intercalibration

As for the NDVI, first on the basis of simulated sensors the WDVI differences between the sensor pairs were analysed (Tab. 4.7). Highest differences due to different RSR functions were again found between SPOT 5 and QuickBird, followed by the sensor pair Landsat 5 TM and QuickBird. Smallest differences were assessed between QuickBird and Aster. Intercalibration studies for the NDVI have shown that using a sixth polynomial order modelling approach reduces the differences significantly. Thus, this approach was also chosen for the WDVI. The sensor differences introduced by different RSR functions decreased significantly (Tab. 4.8), now being below 1% for all assessed sensor pairs. Results are comparable with the findings of the NDVI intercalibration.

Parameter retrieval using the intercalibrated WDVI and the CLAIR model

In Fig. 4.6 the modelled versus observed parameters are displayed using the CLAIR model with the intercalibrated WDVI. For the fresh biomass the scattering is rather large in the mid range biomass between 4 to 5kg/m^2. For the lower and higher range the fit between modelled and observed fresh biomass is satisfactory.
For the dry biomass retrieval findings were the poorest compared to the other biophysical parameters. For early phenological stages the scattering around the 1:1 line is not high, but the R^2 is not as significant as for the other parameters. After the phenological stage heading the scattering increases significantly and the dry biomass is underestimated, even so the R^2 did not decrease that much.

Table 4.7: MIN, MAX and MEAN WDVI differences and percentage difference between the simulated sensors. Percentage differences were calculated considering the WDVI mean of the corresponding sensor and the mean WDVI difference of the respective sensor pair.

sensor pairs	Min WDVI differences	Max WDVI differences	Mean WDVI differences	Differences (%)
SPOT5-Aster	-772	786	123	5.11
SPOT5-Landsat 5TM	-484	828	36	1.52
SPOT5-QuickBird	-676	677	187	7.80
Landsat 5TM-Aster	-511	667	86	3.66
Landsat 5TM-QuickBird	-389	602	151	6.38
QuickBird-Aster	-230	147	-64	-2.91

Concerning the plant water content R^2 are rather high for all phenological stages. For the early stages, the scattering is very low with only some bolters at high plant water content around 4kg/m^2. For later phenological stages R^2 decreased a little, as well as the overall scattering. During these late phenological stages generally an underestimation of the plant water content can be observed.
Results modelling the LAI were also especially for early phenological stages very good featuring very low scattering, whereby it increased slightly for higher LAI values. For late phenological stages there was an underestimation in the higher LAI range obvious.
Generally, results using this approach outperformed the other introduced multispectral models, concerning the scattering and the R^2 for all parameters. Only modelling the dry biomass during late phenological stages was critically. For all other times and biophysical parameters assessed one can conclude that the CLAIR model, actually developed for LAI, can also be used, due to the fact that the derived parameters are all highly correlated with each other. A saturation of the signal could not clearly be assessed, even so the R^2 decreased slightly for higher biomass values and there was a slight underestimation of the derived parameters, the effect was not that distinctive. When comparing this findings with the literature (Clevers, 1988, 1989) it can be recorded that they are in agreement with other studies, maybe the results feature slightly more scattering. This is probably due to the fact that different multispectral sensors (still featuring differences after the intercalibration due to other sensor characteristics), two research years and altogether six different winter wheat fields were assessed.

On the basis of the mapped fields from the ground truth campaign the models were finally inverted for the whole research region and averaged for each mapped field. The ERS-2 data series (section 4.2.4) was used to extract the heading/flowering dates of the individual fields. In Fig. 4.7 exemplarily for the fresh biomass four different dates are shown, two were before heading/flowering (05/04/2006 and 05/11/2006) and two after heading/flowering (06/24/2006 and 14/07/2006). What becomes obvious is the dynamic development of the fresh biomass.

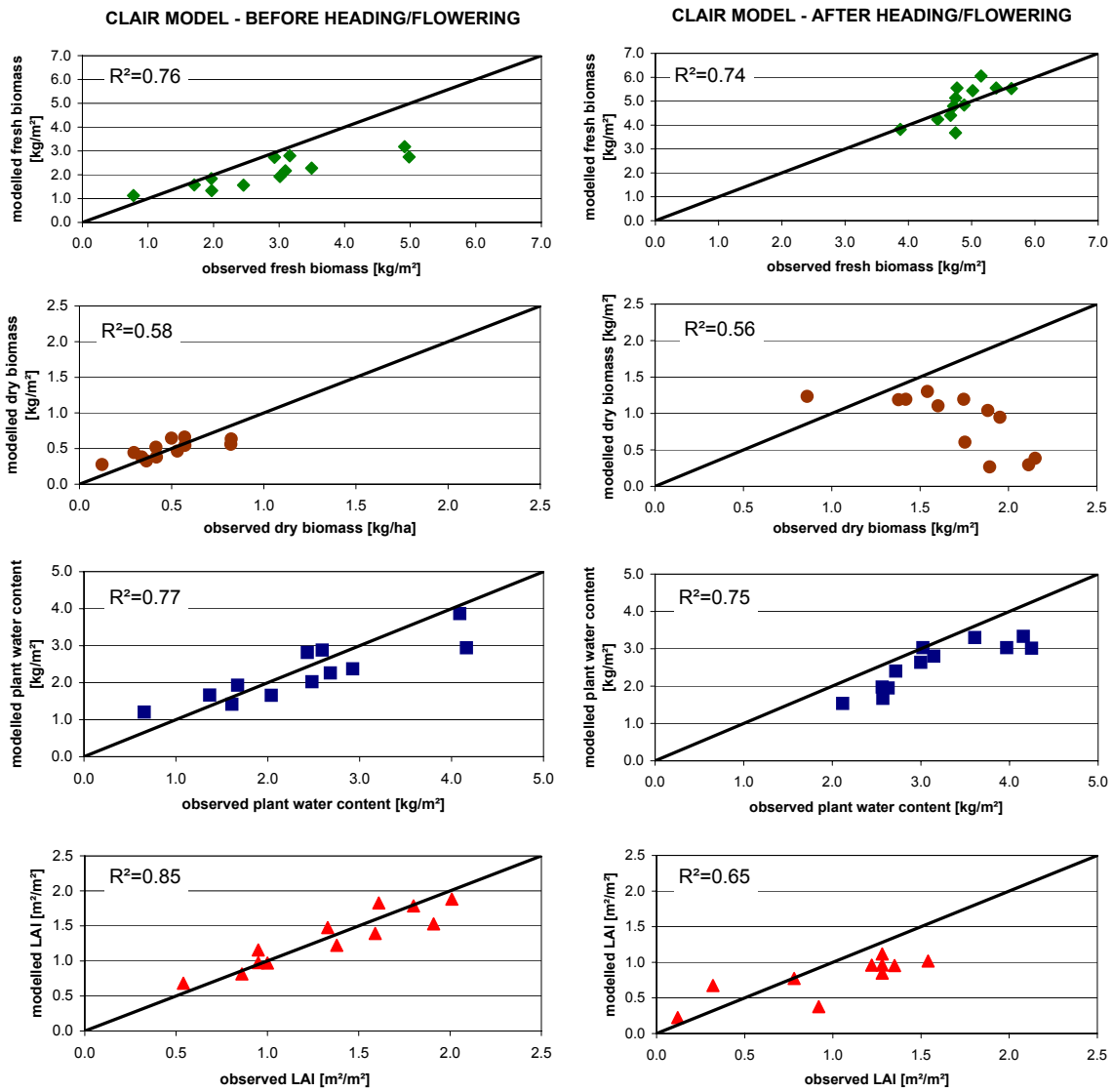

Figure 4.6: Modelled and observed biophysical parameters, divided into before and after heading stage.

Table 4.8: MIN, MAX, MEAN WDVI differences and % difference between the original WDVI imagery and the cross-calibrated image sextic order polynomial modelling.

Differences were taken between the original sensor (org) and the cross-calibrated sensor (cross), which is still named after its origin.

org-cross	MIN	MAX	MEAN	%
SPOT 5-ASTER	-3567	676	20	0.86
SPOT 5-LANDSAT 5 TM	-3628	813	10	0.43
SPOT 5-QuickBird	-2020	375	11	0.49
Landsat 5 TM-ASTER	-365	541	9	0.38
Landsat 5 TM-QuickBird	-646	1376	-0.806	-0.03
Landsat 5 TM-SPOT 5	-707	3018	-11	-0.49
QuickBird-ASTER	-2324	278	5	0.26
QuickBird-Landsat 5 TM	-555	711	1	0.09
QuickBird-SPOT 5	-315	1344	-10	-0.49
ASTER-Landsat 5 TM	-676	309	-10	-0.45
ASTER-QuickBird	-325	882	-9	-0.42
ASTER-SPOT 5	-518	1382	-22	-0.98

For the first acquisition data values are rather low around 1.5kg/m^2. Within one week (second acquisition) the values increase rapidly up to 2.5-3kg/m^2. This rapid increase of around 1kg/m^2, was also measured in the ground truth campaign of the monitored fields. Thus, it is probably due to a real biomass increase and not due to modelling biases. Until end of June the steady biomass increase goes on, up to values of around 6kg/m^2, these high values were also observed within the ground truth campaign. After this date the ripening of the crops start, thus the fresh biomass values decrease again due to the process of drying.

4.1.4 Conclusions

Within this section the overall question:

Is it possible to derive the values of important crop variables from various remote sensing data?

was assessed, whereby in detail it was tried:

- to intercalibrate vegetation indices of different multispectral sensors in order to correct the biases, introduced by differences in the relative spectral response functions
- to evaluate the performance of the CLAIR model for the retrieval of different biophysical parameters
- to invert the used models for regional biophysical parameter extraction

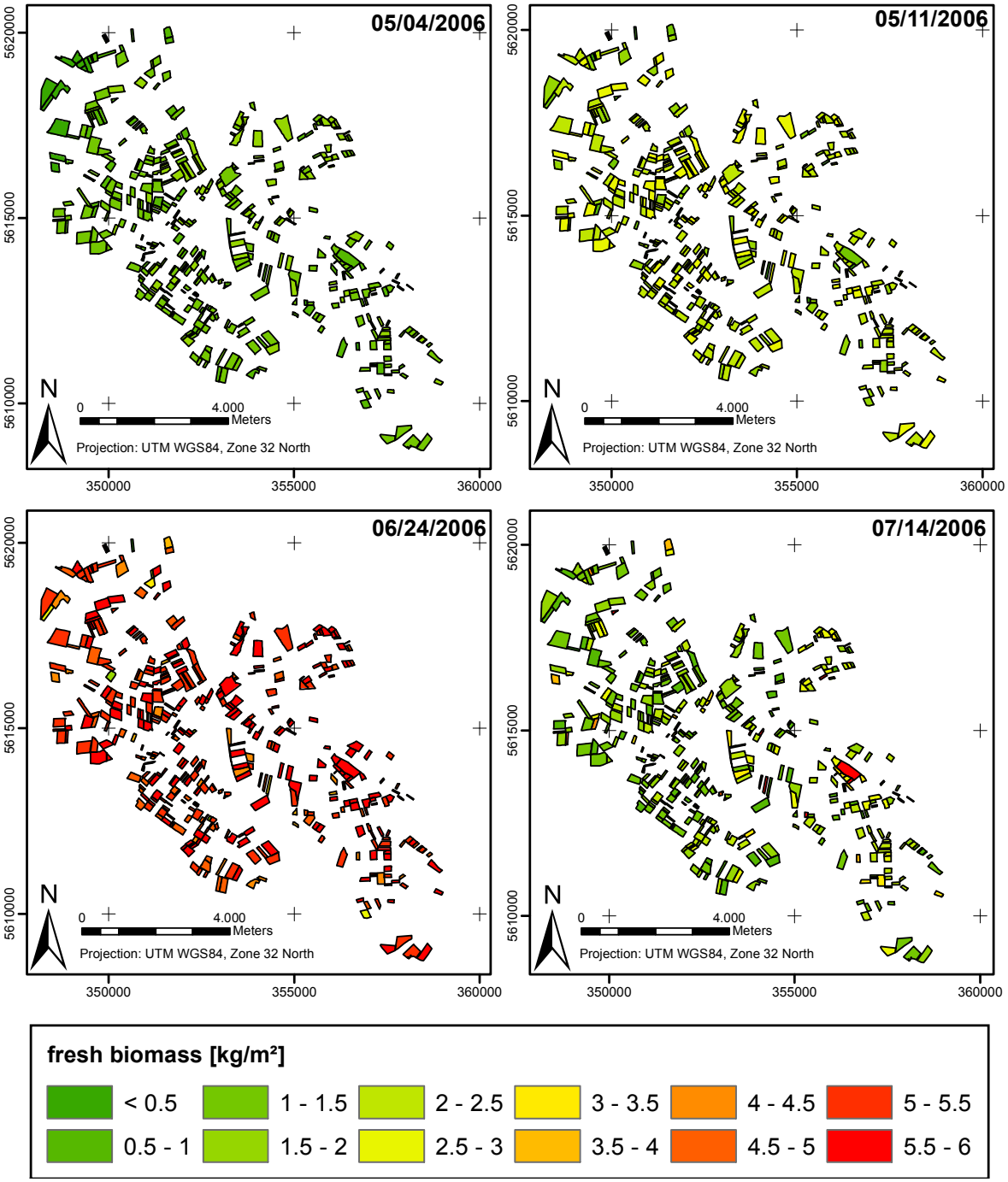

Figure 4.7: Modelled fresh biomass for four different images, using the CLAIR model and the intercalibrated WDVI.

Concerning the overall question, it was possible with all applied approaches to derive the desired parameters, fresh and dry biomass, plant water content and LAI from the multispectral time-series. Generally the different results showed that when using different multispectral data for the derivation of quantitative values an intercalibration of the VIs from the different sensors is appropriate. The NDVI differences decreased rapidly, e.g., between SPOT 5 and QuickBird from 7.087% to 0.152% or 0.161% (depending on the used target sensor), improving the extraction accuracy, e.g., in the case of NDVI retrieval and fresh biomass from R^2 0.67 to 0.73.
In a next step the performance of the CLAIR model was assessed, whereby it was also tried to transfer the model to other biophysical parameters. After the WDVI intercalibration and when considering different phenological stages all biophysical parameters could be derived with a satisfactory accuracy. Findings for dry biomass were overall the poorest, especially after heading/flowering. For the fresh biomass and the plant water content the CLAIR model, even outperformed the empirical models evaluated before. Findings for the parameter LAI, however, were still the soundest, whereby the agreement between modelled and observed was highest before the phenological stage heading.
By the use of this rather simple semi-empirical model it was also possible to regionally extract the desired parameters by simple model inversion.

4.2 Biophysical parameter retrieval using SAR data

In this section also the overall question was again:

Is it possible to derive the values of important crop variables from various remote sensing data?

now assessed for the SAR data, whereby in detail it was tried:

- to analyse the potential of different polarizations and incidence angles for the retrieval of different biophysical parameters (section 4.2.1)
- to analyse the performance of the Water Cloud model and multiple regressions using different SAR data and the additional variable *vegetation roughness* (section 4.2.2 and 4.2.3)
- to extract important phenological events such as heading/flowering from remote sensing information for the adjustment of the retrieval models and the simulation model (section 4.2.4)
- to invert the used models for regional biophysical parameter extraction (section 4.2.2 and 4.2.3)

Like for the multispectral sensors average intensities were calculated for all analysed fields and the parcel boundaries were buffered with 30m zones.
Another processing step was the transformation of the linear backscattering coefficient into decibels:

$$\sigma^0_{(dB)} = 10 \times log_{10}\sigma^0 \tag{4.6}$$

However, for some mathematical calculations, e.g., multiple regression analysis, linear values are more suitable, thus only for the calculations using the Water Cloud Model and for displaying purposes linear values are transformed into decibels.
As an additional information the local incident angle was also calculated by use of the DEM and the respective acquisition angle ranging from near to far range. This information is essential for each field and acquisition date and is especially important when using the ASAR data, due to the differences in the acquisition geometry.

4.2.1 Empirical regressions

In a first step the general relationship between C-band backscatter measurements and biophysical parameters of winter wheat were analysed by using simple empirical regressions for all acquired data in the years 2005 and 2006. This was done by just plotting the backscatter coefficients against the biophysical parameter and assessing the relationship without any further independent validation. No additional information like incidence angle or soil characteristics were considered. Of interest were the performances of the different polarizations and incidence angles for the biophysical parameter estimation.

VV-polarization

In Fig. 4.8 the relationship using a logarithmic regression between the backscatter coefficient [ERS-2] and the different biophysical parameters is displayed. Obvious is that the agreement is only valid for the fresh biomass and the water content. For these two the coefficients of determination (R^2) were significant, but the scattering was rather large in the mid value range. Generally no clear trend of under- or overestimation was obvious. When considering exemplarily the fresh biomass of the two research years individually (Fig. 4.9), slight differences between the two years become obvious. For the year 2005 the coefficient of determination is not as high, however the scattering seems to be less, especially concerning the biomass range from 3 to 5kg/m^2. This might be due to different meteorological conditions or different management practices. Overall there seems to be no major disagreement in the relationship between the two years. In summary concerning the inter-year comparison, the findings seem to remain valid in time. This is important, because due to the small number of data pairs, findings become more stable when looking at both years together. For modelling the leave-one-out cross-validation strategy was chosen, always neglecting one data pair of the time-series 2005 and 2006.

HH and HV-polarizations and ratio

In Fig. 4.10 the relationship between the HH, HV and the ratio (HH/HV) with the different biophysical parameters is shown. A third order polynomial described the patterns best. Only for the fresh biomass the ratio combining the HH and HV polarization outperformed the others, concerning R^2 and the scattering. For the other parameters the HH polarization showed the strongest relationship. For the dry biomass retrieval, the highest R^2 was found, but also the scattering seemed to be more present than for the other parameters. Due to the different incidence angles in a second analysis, different acquisition geometries were taken into account.

When comparing the findings for the different polarizations it becomes appearend that the

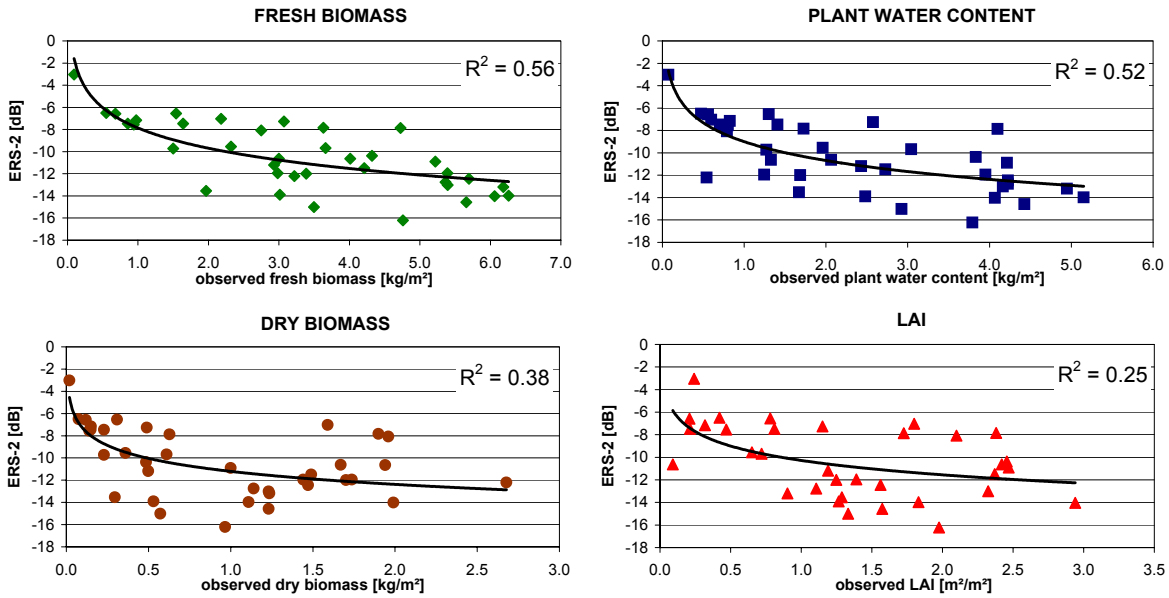

Figure 4.8: Logarithmic relationship between the ERS-2 backscatter and the observed biophysical parameters.

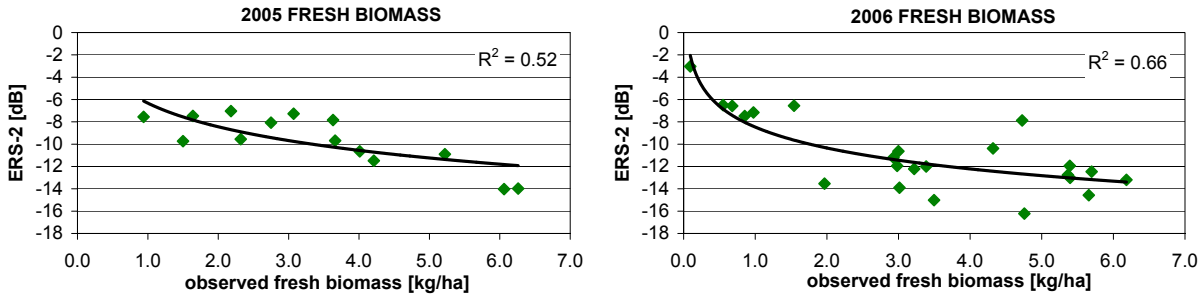

Figure 4.9: Logarithmic relationship between the ERS-2 backscatter and fresh biomass [kg/m^2] for the two research years individually.

HH polarization, or for the fresh biomass the ratio outperform the results achieved for the VV polarization. In order to statistically validate these findings, in a next step multiple regression analysis and actual modelling will be done. Analysis combining the HH and VV polarization could not be assessed, due to the rather large time gaps between the acquisitions of ERS-2 and ASAR data.

Different ASAR swaths

Exemplarily for the biophysical parameter fresh biomass and the relationship between the different ASAR polarizations, under consideration of the two different used swaths (Fig. 4.11), are studied and compared to the above findings (Fig. 4.10). What becomes obvious is that the relationship between the backscatter and the fresh biomass increases when looking at the swath separately. Especially swath 6 features a strong relationship with the biophysical parameters, which is reasonable due to the fact that the soil signal, when using this acquisition geometry is rather low and thus, the proportion of the vegetation signal is higher.
Due to the few ASAR acquisitions and the fact that in the next analysis steps the local incidence angle will be considered additionally, the two different swath will be analysed together.

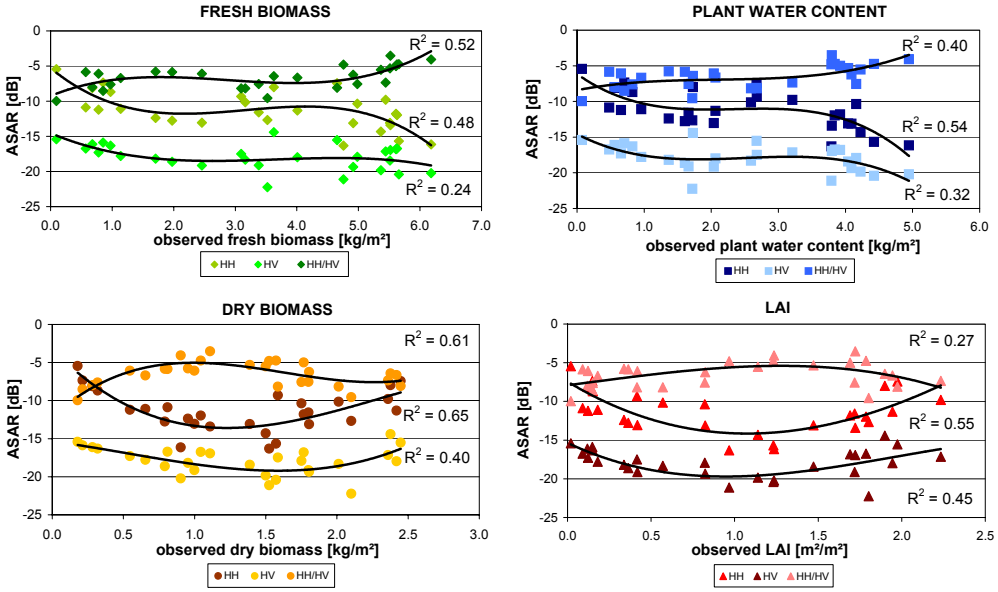

Figure 4.10: Polynomial relationship between the ASAR backscatter and the observed biophysical parameters.

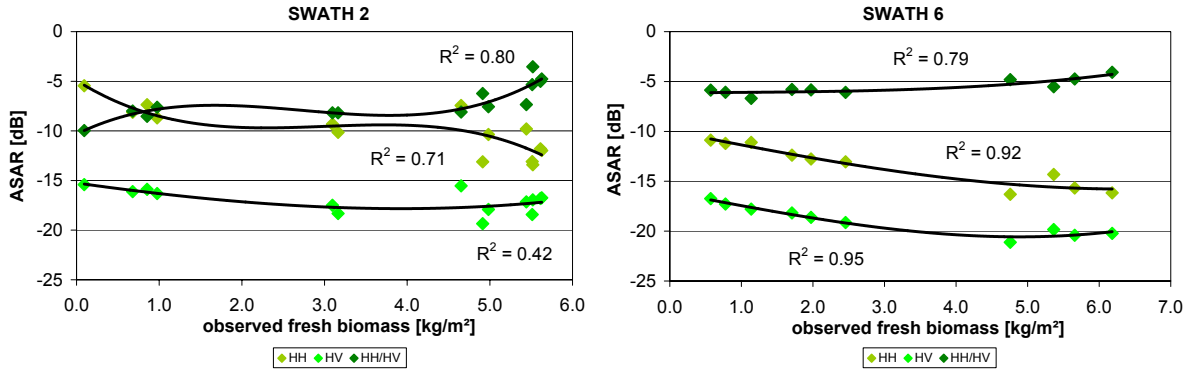

Figure 4.11: Polynomial relationship between the two used ASAR swaths and the observed fresh biomass.

4.2.2 Multiple linear regression analysis

The above findings and the literature show that the backscatter signal is not only influenced by the biophysical parameter, but also by sensor characteristics, e.g., incidence angle and other target characteristics, e.g., soil moisture. Therefore, in a next step modelling and the inversion of the models was done by considering also parameters like soil moisture, represented here by the precipitation sum over five days before the overflight, soil roughness, vegetation roughness and the actual local incidence angle. This was done by taking into account all available data sets, using the leave-one-out cross validation method, based on the term:

$$biophysical\ parameter = A + B \times \sigma^0 + C \times lia + D \times vr + E \times nd + F \times sr \qquad (4.7)$$

with σ^0 the backscattering coefficient [linear], lia the local incidence angle [radian], vr the vegetation roughness [cm], nd the precipitation sum [mm], sr the soil roughness [cm] and the unknown coefficients A to F. The actually fitted coefficients of the respective multiple

regression models are listed in the chapter **ATTACHMENT** (A, Tab. A.2, Tab. A.3, Tab. A.4 and Tab. A.5).
The parameter vegetation roughness, as described in the chapter **DATA** (3) perfectly describes the crop development, firstly increasing until the end of flowering and afterwards with advanced ripening decreasing again, due to crop bending and expending ears. It is correlated with the biophysical parameters to be extracted and might also influence the backscattering coefficients, due to their sensitivity to roughness aspects. For the acquisition times of ERS-2 the R^2 between the respective biophysical parameter and the vegetation roughness were: FM=0.57, DM=0.45, LAI=0.33, PWC=0.36. For the acquisition date of the ASAR sensor a stronger relationship was found with R^2 of: FM=0.65, DM=0.56, LAI=0.28 and PWC=0.5. The strongest impact of the parameter vegetation roughness should be on the retrieval of the fresh biomass using ASAR or ERS-2 data.

Firstly multiple regression analysis are done for all parameters concerning the ERS-2 (VV polarization) time-series, shown in Fig. 4.12. The figure displays the modelled versus the observed parameters together with the coefficient of determination (R^2) for the independent validation process. For the fresh biomass the modelled and observed values are in very good agreement with each other, featuring a very high R^2 with 0.71. The rms error between the estimated fresh biomass and in-situ measurements is at approximately 1.6kg/m^2. The scattering around the 1:1 line increases slightly with increasing biomass values (>3 kg/m^2). Findings for the plant water content are also reasonable, featuring a R^2 of 0.55 (rmse= 1.3kg/m^2). The scattering is quite high for low values in contrast to the fresh biomass. For the parameters dry biomass and LAI the coefficients of determination are not significant with values below 0.5 and rather high divergence of the values around the 1:1 line. For low LAI values there is a trend of value overestimation and for high LAI values an underestimation obvious.

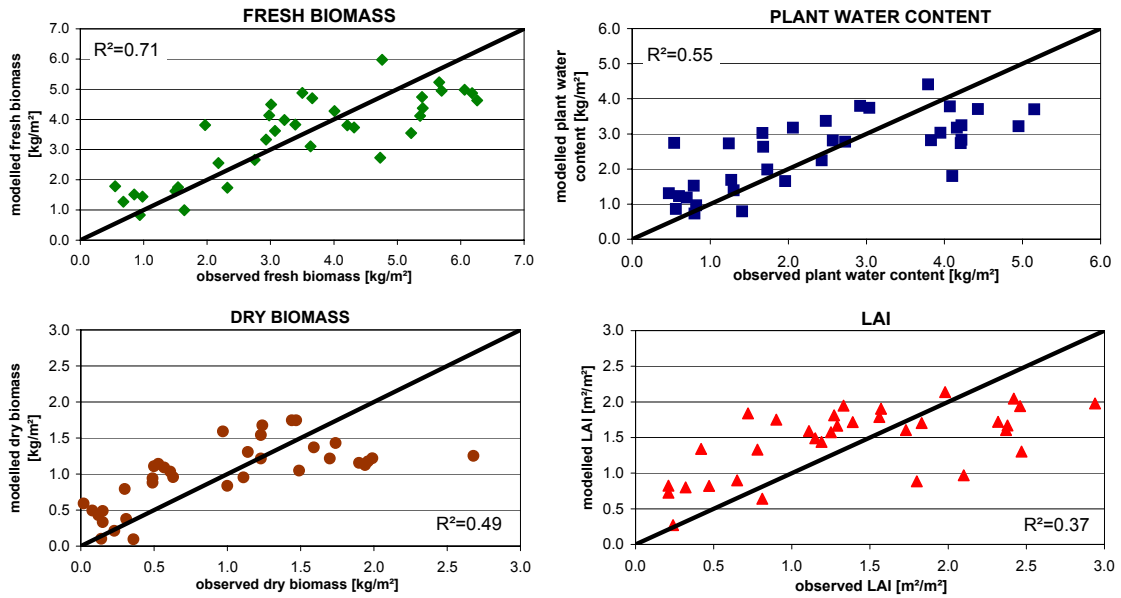

Figure 4.12: Measured versus observed biophysical parameters, modelled by multiple regression using ERS-2 data (VV) from 2005 and 2006.

In Fig. 4.13 in contrast to the findings for the VV-polarization the results for biophysical parameter modelling using HH, HV and the ratio are displayed. Concerning the fresh biomass,

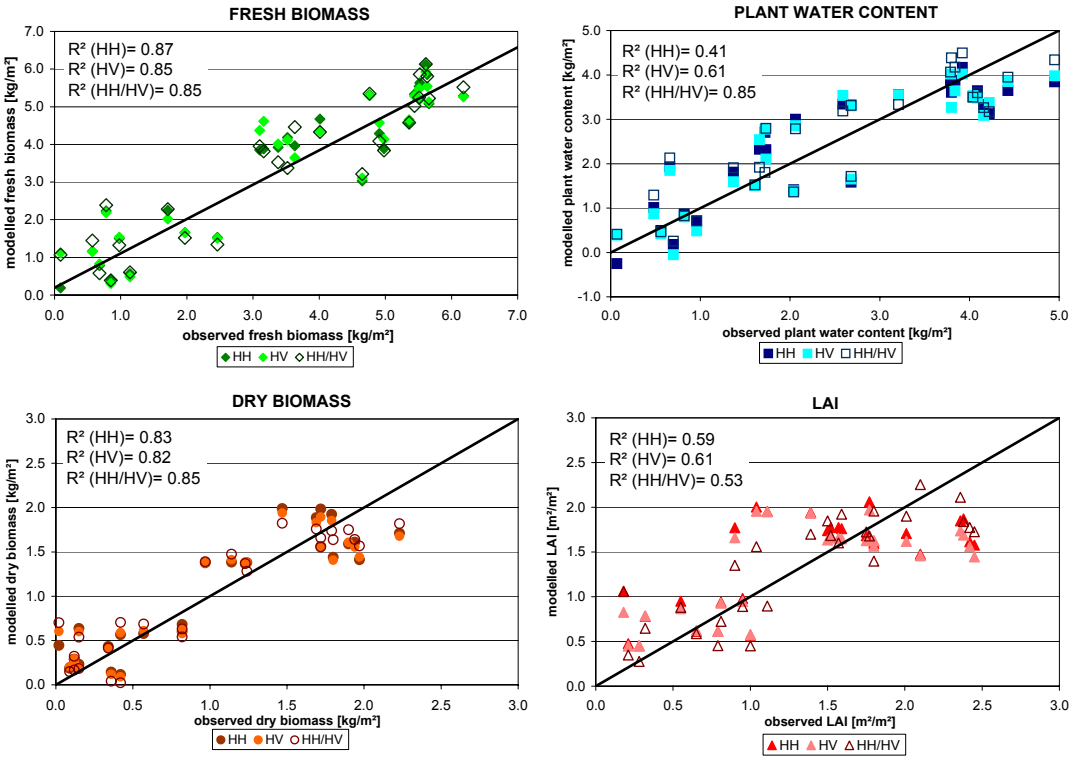

Figure 4.13: Measured versus observed biophysical parameters, modelled by multiple regression using ASAR data (HH, HV and ratio) from 2005 and 2006.

overall the soundest results were achieved, when using the HH polarization ($R^2 = 0.87, rmse = 1.6kg/m^2$). Also the performances when using HV polarization or the ratio are very good and scattering is rather low for all three models. The biggest scattering can be observed around the fresh biomass range from 2 -3 and 4.5 - 5kg/m^2. Concerning the plant water content the R^2 of the three models were quite different from each other. The highest relationship was found when using the ratio ($R^2 = 0.85, rmse = 1.2kg/m^2$) and the lowest when using the HH polarization ($R^2 = 0.41, rmse = 1.3kg/m^2$). The overall scattering was not that high and a trend of over or underestimation could not be clearly detected. For the parameter dry biomass the ratio ($R^2 = 0.85, rmse = 0.6kg/m^2$) outperformed the models using a single polarization. For larger dry biomass values a slight tendency of underestimation can be detected. Concerning the parameter LAI modelling with the HV polarization had the highest R^2 (0.61)($rmse = 0.6$). When comparing the scattering behaviour with the other parameters, it is rather large in regard to the small data range, but also not featuring a clear trend. For the parameters FM and LAI these findings are slightly different to the previous first relationship analysis. This is probably due to the fact that for the first time now an independent validation was performed and also other features are considered. For the DM and the PWC the same polarization (HH) as in the first general overall trend analysis still performed best. Generally, it was not possible to define a here studied ASAR polarization, performing sufficently for all considered biophysical parameters.

When comparing the results achieved using the ASAR sensor (HH, HV, ratio) with the ERS-2 sensor (VV) it becomes apparent that the models using the ASAR sensor outperformed (except for HH polarization for plant water content) the ERS-2 sensor. The results are compared with the literature (Dente et al., 2007; Mattia et al., 2003) appropriate and within a reasonable range.

Due to the fact that the results will be later used for regional biophysical parameter extraction,

the parameters vegetation and soil roughness are not that appropriate to use. For the regional extraction also the parameter soil moisture could not be used and therefore the soil moisture impact was represented, as also already done in the previous analysis, by precipitation sums. For regional applications it is important to only use model input parameters, which can be obtained without a complex ground truth campaign of the whole region. In a next step multiple regression analysis was done only using the term:

$$biophysical\ parameter = A + B \times \sigma^0 + C \times lia + E \times nd \tag{4.8}$$

with σ^0 the backscattering coefficient [linear], lia the local incidence angle [radian], nd the precipitation sum [mm] and the unknown coefficients A, B, C and E.

Fig. 4.14 displays the observed versus the modelled biophysical parameters considering the ERS-2 sensor (VV). Like for the results in Fig. 4.12 only for the biophysical parameters fresh biomass and plant water content significant retrieval results, with $R^2 > 0.5$ could be achieved. Whereby, for both the scattering has increased a little ($rmse = 1.7kg/m^2$ (FM), $rmse = 1.4kg/m^2$ (PWC)). There is again for both a trend of value overestimation within the low value range and of underestimation within the high value range obvious. Generally, it seems that the roughness parameters have a higher influence, when retrieving the fresh biomass, as assumed previously, due to the high correlation of the vegetation roughness and the fresh biomass. In Fig. 4.15 achievements using the ASAR sensor are displayed. For the fresh biomass, results

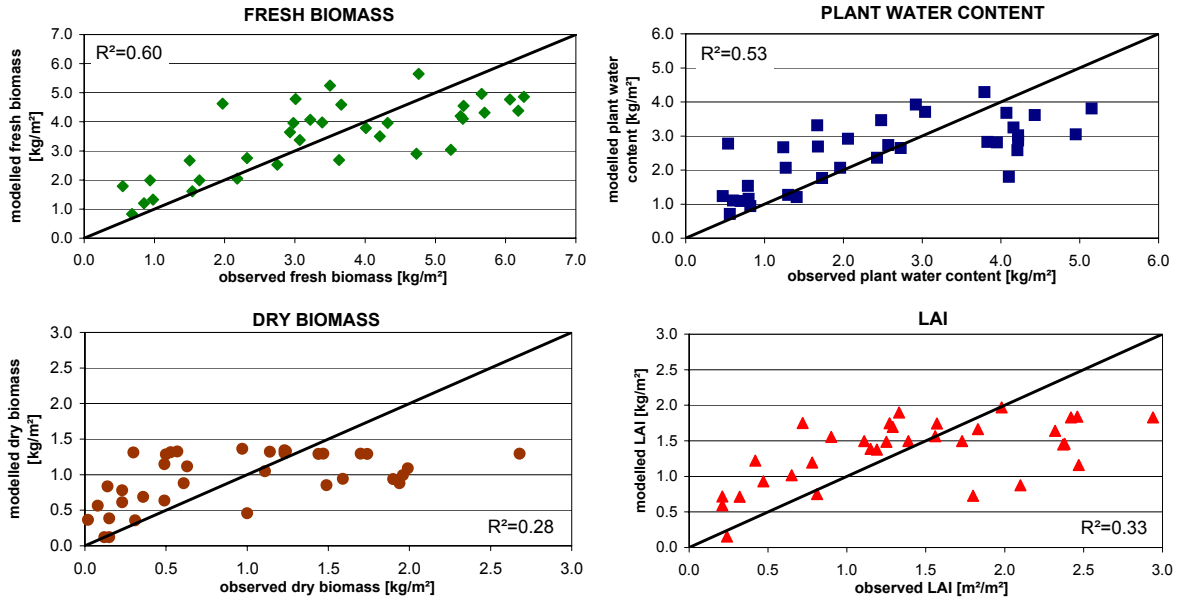

Figure 4.14: Measured versus observed biophysical parameters, modelled by multiple regression (without considering vegetation and soil roughness) using ERS-2 data from 2005 and 2006.

have decreased for all polarizations, compared to the multiple regressions shown in Fig. 4.13. Especially, for the ratio the scattering has increased significantly (from rmse 1.7 to 1.9kg/m^2), which becomes very obvious in the high biomass range. In contrast to the previous results now modelling using the HV polarization outperforms the others concerning R^2 (0.83) and the scattering behaviour (rmse=1.7kg/m^2), especially obvious in the biomass range above 4.5kg/m^2. It seems that the roughness parameters, especially have an influence on the HH polarization and the ratio. As expected previously due to the high correlation between the fresh biomass and the

vegetation roughness the modelling results have decreased. For the plant water content there is no clear relationship any more for all studied polarizations. The coefficients of determination are all below 0.5 and the scattering increased drastically over the whole range. For this parameter the backscatter influence by vegetation and soil roughness is rather high. For the dry biomass the coefficients of determination did not decrease as much. The scattering, regarding modelling with the HH polarization increased (from rmse=0.7 to 0.8kg/m^2), especially in the high biomass range. For the HV and the ratio the scattering and R^2 nearly stayed the same, here the roughness informations have the largest impact on the HH polarization. When using the ratio the modelling results slightly outperform the HV polarization concerning the scattering (rmse: HV= 0.69, HH/HV= 0.67). Concerning the biophysical parameter LAI, only for the HV polarization significant results in regard to R^2 and the scattering behaviour (rmse=0.6kg/m^2) were achieved. The largest scattering can be found around LAI values between 1 and 1.5. Even so for the ratio, R^2 is above 0.5, the scattering is rather high. For this parameter again roughness informations had the highest impact, when modelling with the HH polarization.

In this analysis it is possible to define a suitable polarization for all biophysical parameters

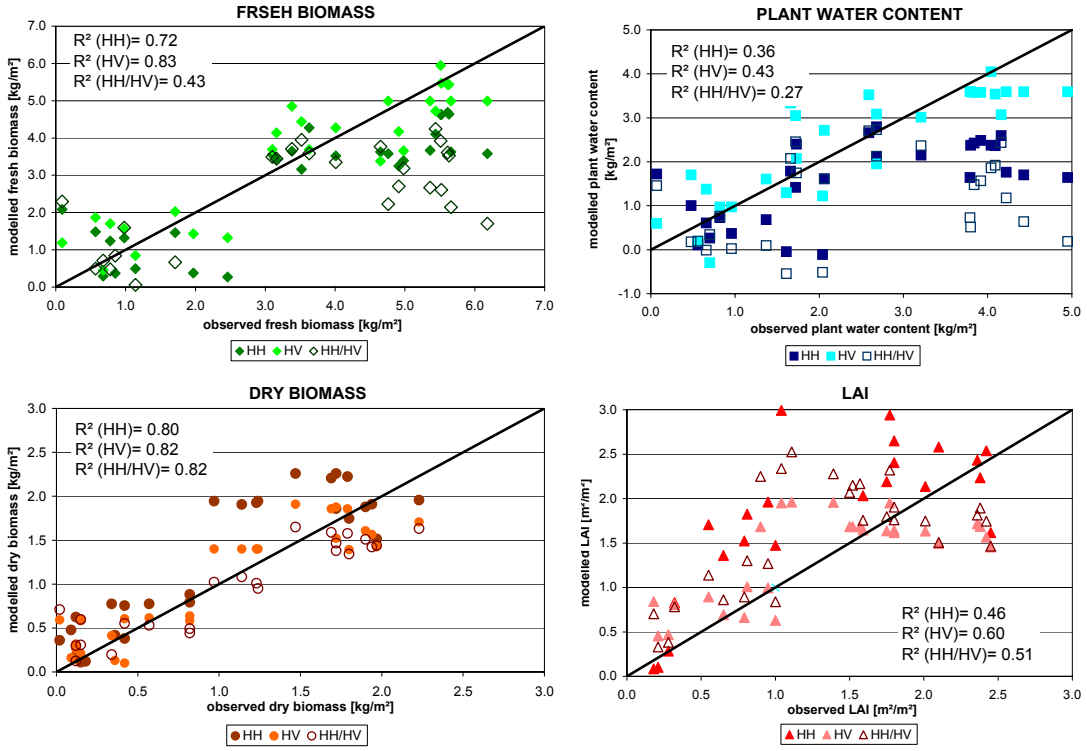

Figure 4.15: Measured versus observed biophysical parameters, modelled by multiple regression (without considering vegetation and soil roughness) using ASAR data (HH, HV and ratio) from 2005 and 2006.

together. The HV polarization outperforms all other polarizations, which stands in contrast to the previous findings (except for the DM).

Overall, in regard to all derived parameters and polarizations (Fig. 4.14 and Fig. 4.15) the analysed ASAR polarizations exceeded the ERS-2 polarizations, except for the parameter plant water content. For all other parameters using the ASAR sensor gave more reliable results. In agreement to the findings from the section **Empirical regressions** (section 4.2.1) the findings underline the fact that the used ASAR polarizations outperform the VV polarization from ERS-2, for the retrieval of nearly all analysed biophysical parameters (except plant water content).

The assumption that the HH polarization for fresh biomass and for the others the ratio would perform best could not be stressed. Instead the HV polarization surpasses the others. Only in the case of the dry biomass, the assessed scattering is slightly lower, when modelling with the ratio.

It might be possible to actually improve the results by using generalized assumptions of the vegetation and soil roughness over the vegetation period. For the vegetation roughness one could presume increasing values until end of flowering and afterwards decreasing values. For the soil roughness one could imply high values during early phenological stages and lower values with increasing plant development. Due to the fact that the precipitation has a decreasing impact on the soil roughness and because with increasing vegetation coverage, the soil impact on the backscattering coefficient decreases as well.

4.2.3 Semi-empirical Water Cloud Model

In a next analysis step the retrieval performance of the semi-empirical Water Cloud Model was analysed considering all available polarizations. Especially, of interest were the results for extracting the plant water content from the used ASAR polarizations and the biophysical parameters dry biomass and LAI from the ERS-2 VV polarization, because the just previously mentioned results were not valid. Firstly retrieval results, when also using the parameter vegetation roughness are discussed and afterwards, the model without considering the roughness parameters. Model calibration and validation, based again on the leave-one-out cross-validation method, was used. The individually fitted coefficients for each model are listed in the chapter **ATTACHMENT** (A, Tab. A.6, Tab. A.7, Tab. A.8, Tab. A.9).

The assessment of the impact, when integrating the vegetation roughness into the semi-empirical Water Cloud Model, the model original introduced in chapter 2.2.3 was extended by $A \times Vege\theta$:

$$\sigma^0 = \frac{-2 \times B \times b_p}{cos\theta + C(\theta) + A \times Vege\theta + D \times m_v} \tag{4.9}$$

with σ^0 expressed in dB, θ the incidence angle [radian], the coefficients A,B,C,D, b_p the respective biophysical parameter, $Vege$ the vegetation roughness in cm and m_v the soil moisture (expresses as precipitation sum). In the literature the impact of the soil roughness is described as incidence depended, therefore the vegetation roughness is formulated as depended on the incidence angle.

In Fig. 4.16 the modelled versus observed biophysical parameters are displayed with the 1:1 line, as a first indicator of the relationship and the respective R^2. Explicit is that for all parameters there was a trend of underestimation in the higher value regions. Only for the parameter fresh biomass a significant $R^2 > 0.5$ was achieved. When comparing the findings with Fig. 4.12, R^2 has decreased while the rmse has increased to 1.7kg/m^2. For all other observed parameters results are even less consistent, featuring rather dominant scattering.

In the next figure (Fig. 4.17) modelling results for the ASAR sensor using the different polarizations are displayed. The results are not really valid, featuring a clear trend of underestimation for all parameters and polarizations. Only for the fresh biomass (HH polarization) and the plant

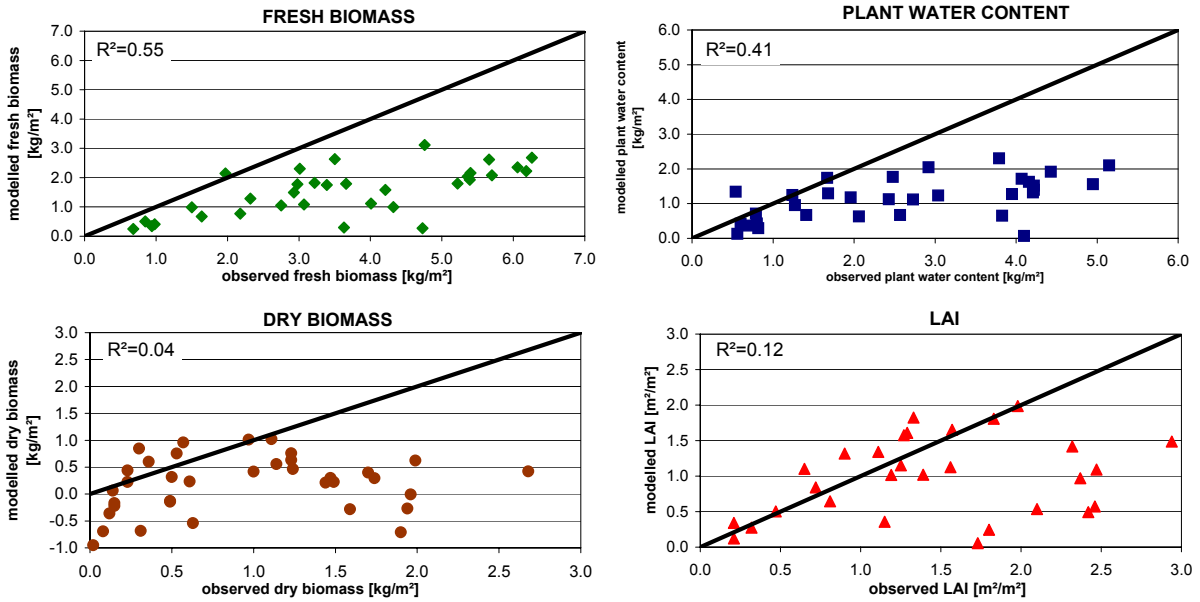

Figure 4.16: Measured versus observed biophysical parameters, modelled with the Water Cloud model extended by the parameter vegetation roughness (ERS-2 data from 2005 and 2006).

water content (HH polarization) significant R^2 could be observed. Compared to Fig. 4.13, results for the fresh biomass and the plant water content have decreased featuring also a high scattering.

Due to the already mentioned problem of the need for the model to be utilized to larger regions and the not very sound results just described, in a next step the performance of the original Water Cloud Model, without considering roughness parameters, was assessed (section 2.2.3).
In Fig. 4.18 the results using the ERS-2 data are displayed. For the parameters fresh biomass and plant water content a clear trend of underestimation within the higher value region can be observed. For the other two parameters, a rather dominant scattering around the 1:1 line, featuring no clear trend, becomes obvious. Compared to Fig. 4.16, there was no clear modelling improvement. Results for the fresh biomass and the plant water content stayed nearly the same and for the other two the scattering increased enormously. Concerning the modelling performances using the ASAR sensor (HH, HV, ratio) displayed in Fig. 4.19, there is again a clear trend of underestimation for the higher value range. Only for the LAI using HH or HV polarization there is actually a distribution around the 1:1 line, featuring a tremendous scattering behaviour. Compared to Fig. 4.17 the modelling performances did really improve, still exhibiting rather high biases for all analysed parameters and polarizations.

Overall, when comparing all modelling approaches using SAR data, simple multiple regression analysis outperformed the semi-empirical modelling. Best results were achieved when also considering roughness features, however by virtue to regional application, it is more reasonable to neglect them.
When using the ERS-2 sensor (VV) only the retrieval of fresh biomass and plant water content delivered significant results. For the ASAR sensor, modelling fresh biomass and LAI using the HV polarization or the dry biomass using the ratio was appropriate.
The use of the Water Cloud Model delivered no valid results, as opposed to literature conclusions, where especially for the plant water content retrieval results were reasonable (Attema & Ulaby (1978); Jongschaap & Schouten (2005); Prevot et al. (1993)).

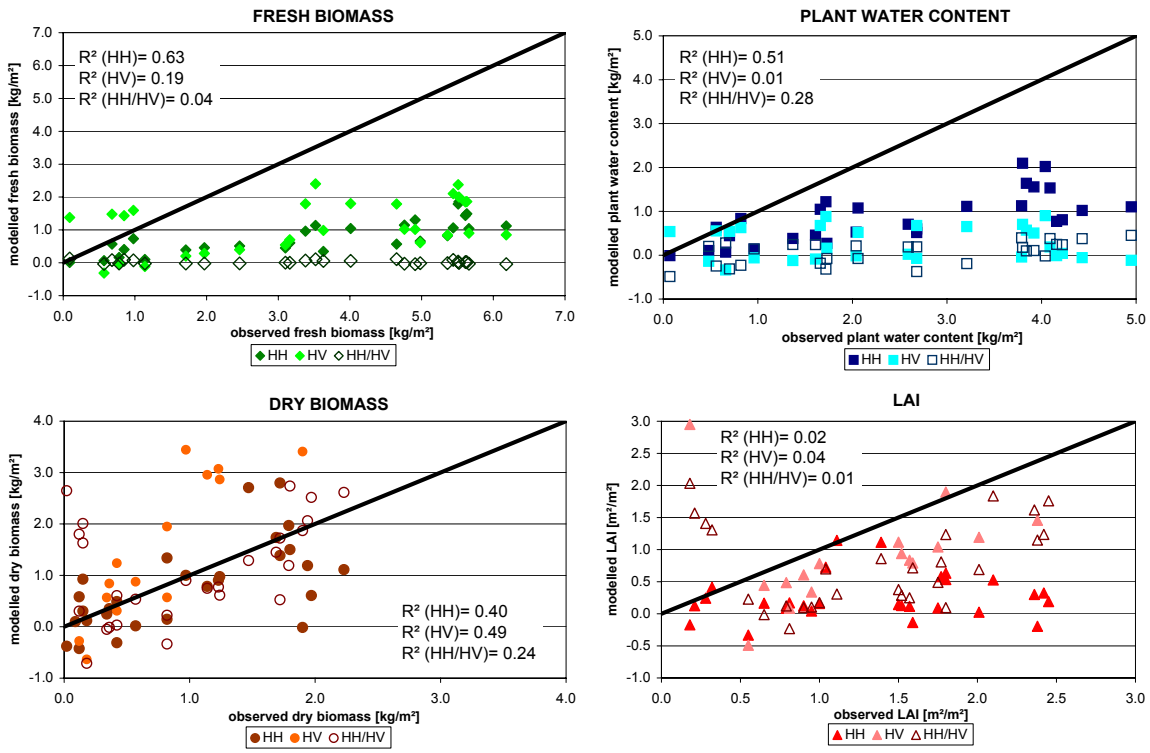

Figure 4.17: Measured versus observed biophysical parameters, modelled with the Water Cloud model extended by the parameter vegetation roughness, using the ASAR data from 2005 and 2006.

On basis of the mapped fields during the ground truth campaign in 2006 the multiple regressions were inverted and applied to all winter wheat fields in the region of Meckenheim. In Fig. 4.20 exemplarily for four images, fresh biomass maps are calculated. During the phenological development fresh biomass increases until ripening. The heterogeneity of the mapped fields is the largest for images taken during maximum LAI (fresh biomass). For the first and last fresh biomass maps shown, the value variety is not as high. In this figure also the trend of over- and underestimation, as shown in Fig. 4.14 becomes obvious. As the multi-temporal fresh biomass dynamics, especially between the last three dates is not as dominant as the monitored ground truth information suggested.

4.2.4 Heading/flowering date

Flowering is a very important and distinctive phenological event in wheat production, marking the start of grain growth in ears at the top of the canopy, it is an essential parameter for crop growth modelling. Ear biomass per unit area increases over time, through the increase in the number of flowering plants, through grain growth from current photosynthesis products and through translocation of carbohydrates from temporary storage organs to the developing grains. The growing ears significantly affect the SAR backscatter signal in VV polarization, through the increasing biomass at the top of the canopy. This modifies the crop geometry and crop moisture distribution, hence attenuating the SAR signal. ERS-2 time-series (late 2004 to late 2006) of winter wheat fields were assessed for the extraction of the phenological stage *heading/flowering*

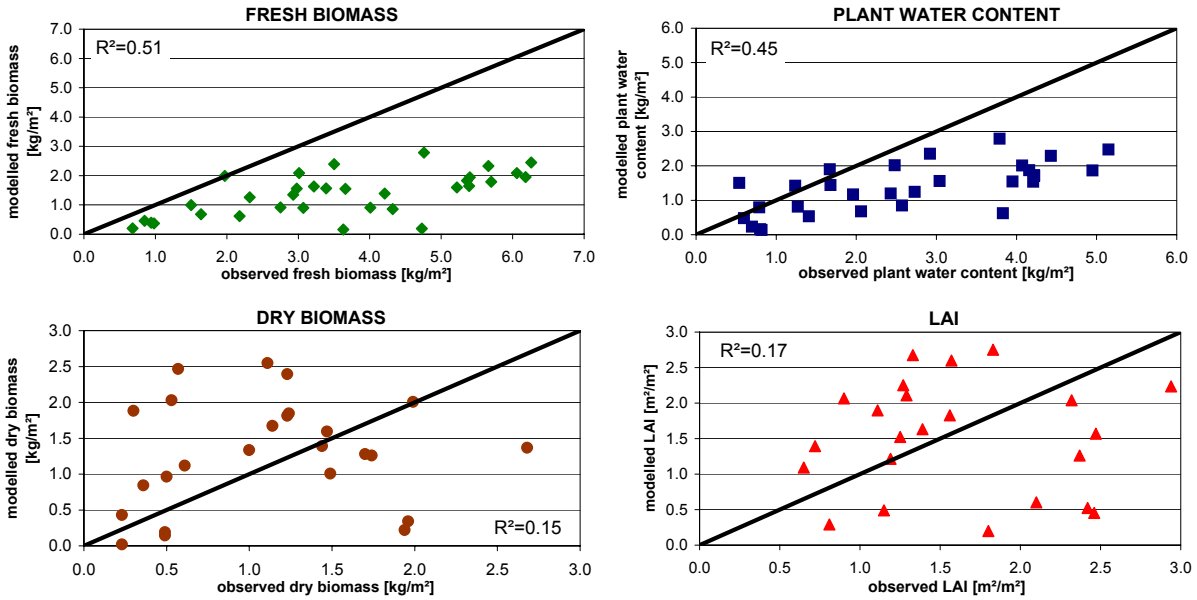

Figure 4.18: Measured versus observed biophysical parameters, modelled with the Water Cloud model using the ERS-2 data from 2005 and 2006.

(Fig. 4.21). During phenological development, both curves start with a local maximum for soil backscattering, to a local minimum, when the soil backscattering is fully suppressed at heading/flowering, continuing to a local maximum at maturity, when the soil characteristics increasingly influence the signal again. The heading/flowering date of the winter wheat crops can be detected by finding the local backscatter minimum (i.e. at maximum attenuation) in the studied ERS-2 time-series. It denotes the maximum water content per unit surface area and hence the heading/flowering date. For the two fields shown in Fig. 4.21 this is the case at the beginning of June. The findings were in a second step extracted to the whole research region. In order to enable the extraction of the exact date, it is important to have a dense time-series, especially during the important phenological development stages.

4.2.5 Conclusions

In this section also the overall question:

Is it possible to derive the values of important crop variables from various remote sensing data?

was assessed, now using the SAR data, in detail it was attempted:

- to analyse the potential of different polarizations and incidence angles for the retrieval of different biophysical parameters
- to analyse the performance of the Water Cloud model and multiple regressions using different SAR data and the additional the variable *vegetation roughness*

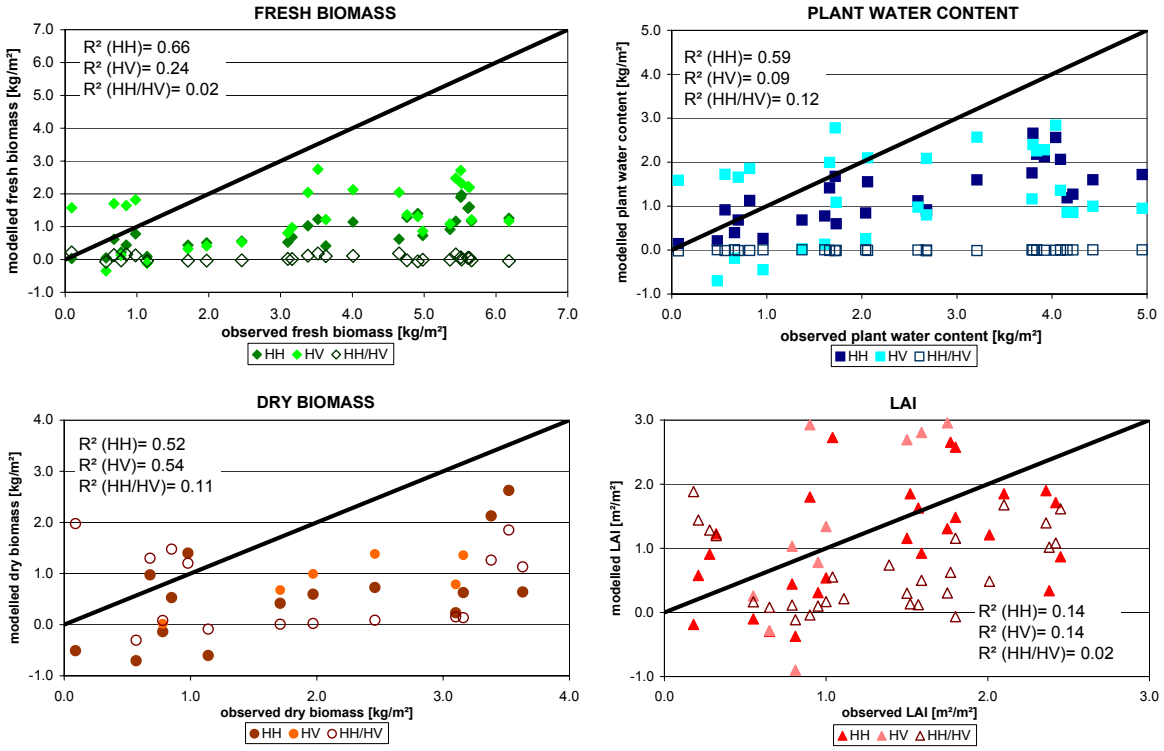

Figure 4.19: Measured versus observed biophysical parameters, modelled with the Water Cloud model using the ASAR data from 2005 and 2006.

- to extract important phenological events such as heading/flowering from remote sensing information for the adjustment of the retrieval models and the simulation model

- to invert the used models for regional biophysical parameter extraction

Considering the potential of the different polarizations and incidence angles for retrieval purposes the findings were contradictory. Satisfactory results for the VV polarization could only be achieved for the retrieval of fresh biomass and the plant water content. For the ASAR sensor modelling fresh biomass and LAI using the HV polarization or the dry biomass using the ratio was appropriate. It was not really possible to make an overall assumption about the most appropriate polarization to use. Concerning the incidence angle, first simple analysis have shown that higher incidence angles, like the ASAR swath 6 outperform lower incidence angles, due to a smaller soil signal amount in the backscattered.
Regarding the performance of empirical and semi-empirical modelling and the additional parameter vegetation roughness, modelling approaches using simple multiple regression analysis outranged the semi-empirical modelling. Best results were achieved when also considering roughness features, due to the need of regional modelling and operational use, it is more appropriate to not consider them. Without considering the roughness features, a reasonable regional retrieval of all desired parameters was achieved. The phenological stages heading/flowering could be derived satisfactorily from the VV time-series. The VV time-series was dense enough during the important and highly dynamic stages of the crop growth period. Findings could also be easily applied for regional extraction of the heading/flowering date.

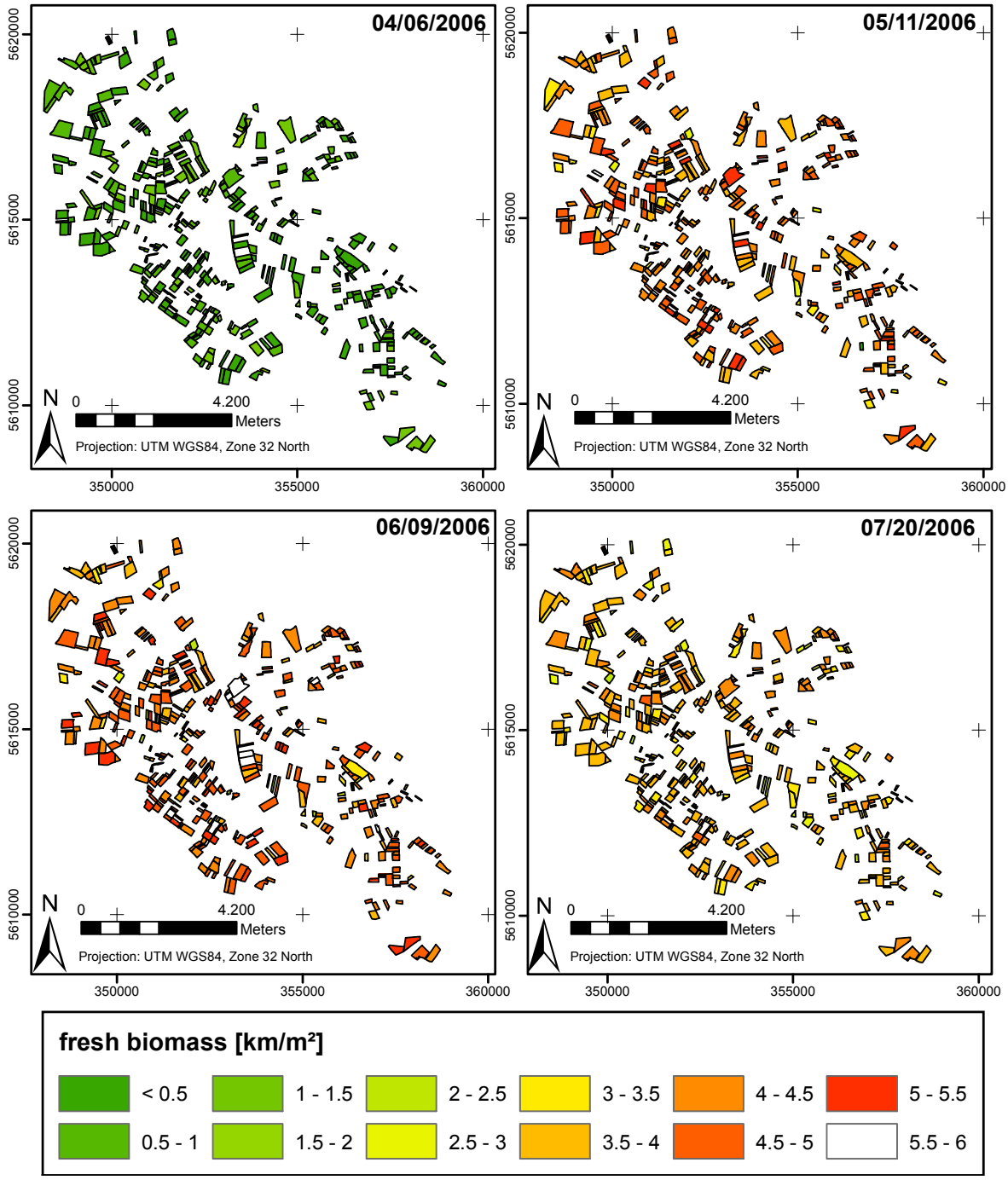

Figure 4.20: Modelled fresh biomass for four different ERS-2 acquisitions, using multiple regression without roughness information.

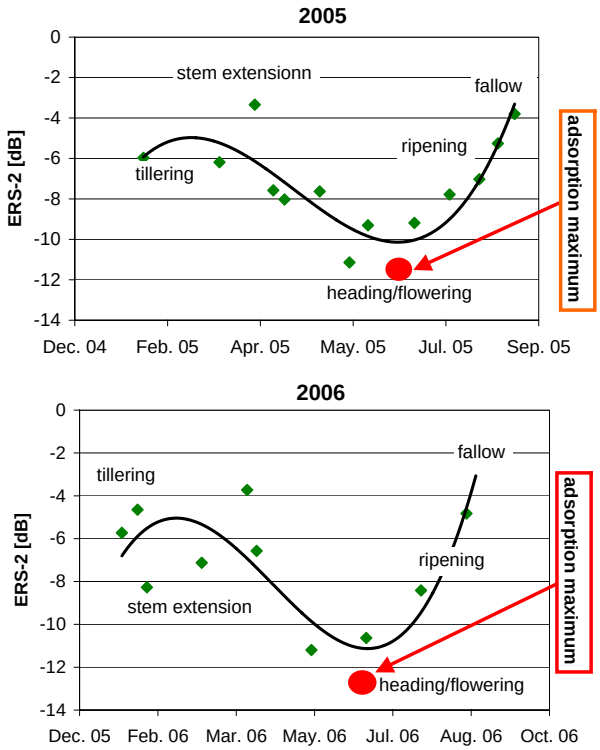

Figure 4.21: Identification of the heading/flowering date using the ERS-2 time-series, exemplarily shown for one winter wheat field within each study year.

4.3 Biophysical parameter derivation using multi-sensoral data

Within the following section the overall question was:

Is it possible to derive the values of important crop variables from various remote sensing data?

assessing in further detail:

- the potential of combining multispectral and SAR data for the derivation of biophysical parameters (section 4.3)
- the inversion of the used models for regional biophysical parameter extraction (section 4.3)

For the combined modelling approach only those data pairs (multispectral and SAR) were considered with acquisition dates within one week. This was done, because otherwise the biophysical field parameters at the focus of the retrieval, would have changed to drastically mean wise. In Tab. 4.9 multispectral and ERS-2 data pairs and in Tab. 4.10 for ASAR and the multispectral data, together with the used ground truth data are listed. As for the previous studies the multispectral vegetation indices used (NDVI, WDVI), were all intercalibrated to the reference sensor Landsat 5 TM. The modelling was done with the leave-one-out cross-validation method.

Table 4.9: Multspectral and ERS-2 data pairs together with the used ground truth date (GT).

multispectral sensor	date	ERS-2 date	GT
ASTER	04/03/05	04/15/05	04/11/05
QuickBird	04/22/05	04/21/05	04/25/05
Landsat 5 TM	05/05/05	05/10/05	05/09/05
Landsat 5 TM	05/28/05	05/06/05	05/30/05
QuickBird	06/20/05	06/24/05	06/20/05
SPOT 4	05/04/06	04/25/06	05/03/06
SPOT 2	05/11/06	05/11/06	05/10/06
SPOT 5	06/24/06	06/15/06	06/21/06
SPOT 2	07/14/06	07/20/06	07/19/06

Table 4.10: Multspectral and ASAR data pairs together with the used ground truth date (GT).

multispectral sensor	date	ASAR date	Swath	GT
ASTER	04/03/05	04/12/05	6	04/11/05
SPOT 4	05/04/06	05/02/06	6	05/03/06
SPOT 5	06/24/06	06/25/06	2	06/28/06

4.3.1 Combined modelling using a simple linear approach

In a first analysis a combined modelling approach using the NDVI and the backscattering coefficients by linear regression was assessed (Brakke et al., 1981), by using the term:

$$b_p = A + B \times NDVI + C \times \sigma^0 \tag{4.10}$$

with b_p the respective biophysical parameter, the intercalibrated vegetation index NDVI, the backscattering coefficient σ^0 and the unknown to be fitted coefficients A,B and C. The fitted coefficients for this approach are listed for all analysed polarizations and biophysical parameters in Tab. A.10 in the chapter **ATTACHMENT** (A).
In Fig. 4.22 the retrieval results for combining the NDVI with the ERS-2 VV backscattering coefficients are displayed, together with the 1:1 line, the R^2 and the rmse for all derived biophysical parameters. Statistically reasonable results could only be achieved for the dry biomass and the plant water content. When comparing the findings with the performances of the multispectral CLAIR model there was a decrease in the accuracy. Compared to the performance of the multiple regression results (VV polarization), there was a slight improvement for the plant water content and the LAI retrieval. Generally combing the two sensors does not make sense in this case, one should rather instead use the multispectral data alone.

In Tab. 4.11 the statistical results for combining the NDVI with the ASAR HH, HV polarization or the ratio are shown. Concerning the fresh biomass valid results could be achieved for the HH polarization and the ratio combinations, both featuring rather low rmse. Combing the NDVI with the ratio actually slightly outperforms findings for the CLAIR model or the multiple regression results, when only using the ratio. The combination also outperforms the HH polarization modelling, concerning the height of the scattering. For this biophysical parameter a combined modelling using the ratio and the NDVI makes sense. In this case actually the ratio

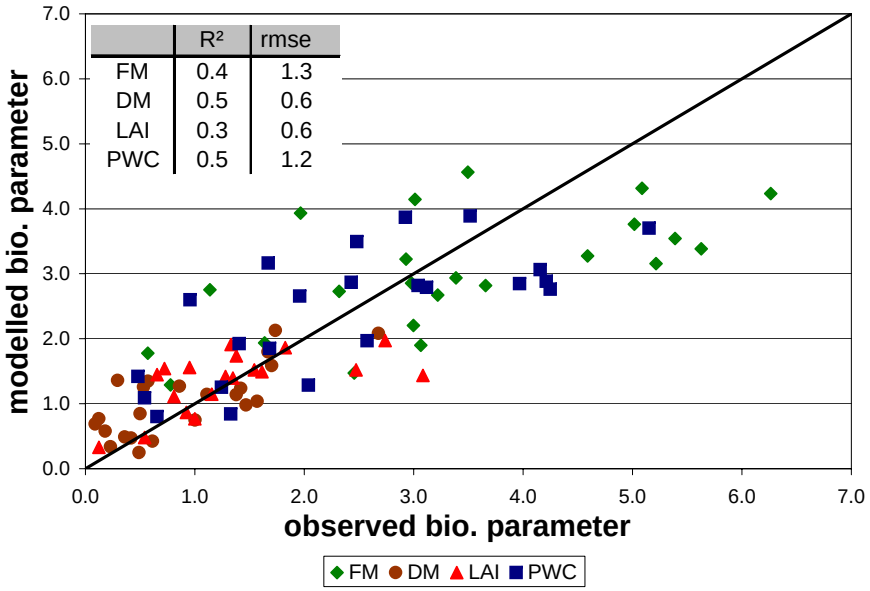

Figure 4.22: Combined biophysical parameter modelling using linear regression, considering the NDVI (intercalibrated) and the ERS-2 backscatter.
FM=fresh biomass [kg/m^2], DM= dry biomass [kg/m^2], LAI [m^2/m^2], PWC= plant water content [kg/m^2]

exceeded the other used ASAR polarizations, whereby for the single sensor case it was the HV polarization.
Concerning the dry biomass, the combination with the ratio clearly surpasses the CLAIR model. In regard to the multiple regressions using the ASAR sensor, even though the R^2 for the non combined modelling is higher, for the ratio the rmse has decreased from 0.6 to 0.38, giving a sounder performance. For the HH or HV polarization the findings were not as valid. Also for the single sensor case modelling using the ratio outperformed the other polarizations.
Concerning the plant water content and the LAI the combined approach outnumbered the single ASAR use for all polarizations. Using the CLAIR model for the retrievals, still gives sounder results, featuring less scattering. Overall, for these parameter it makes no sense to combine the sensors by linear regression.

4.3.2 CLAIR model combined with multiple linear regression for SAR data

As findings for the NDVI and ERS-2 VV combination approach were not as sound, in a next step other fusion approaches were assessed. The fusion was realised by combining the CLAIR model (before and after heading/flowering) with the backscatter coefficient, the local incidence angle of the ERS-2 acquisition and the precipitation sums, were analysed by using the following term:

$$b_p = W_{CLAIR} \times \left[-\frac{1}{\alpha} ln \left(1 - \frac{WDVI}{WDVI_\infty} \right) \right] + W_{ML} \times \left[A + B \times \sigma^0 + C \times lia + E \times nd \right] \tag{4.11}$$

Table 4.11: Combined biophysical parameter modelling using linear regressions, considering the NDVI (intercalibrated) and the ASAR backscatter.

bio. parameter	pol.	R^2	rmse
FM	HH	0.60	0.85
	HV	0.08	0.87
	HH/HV	0.76	0.88
DM	HH	0.35	0.73
	HV	0.41	0.34
	HH/HV	0.78	0.38
PWC	HH	0.53	0.57
	HV	0.01	0.46
	HH/HV	0.59	0.59
LAI	HH	0.63	0.26
	HV	0.64	0.25
	HH/HV	0.66	0.23

bio. parameter= biophysical parameter, pol.=polarization, FM=fresh biomass [kg/ha], DM= dry biomass [kg/ha], LAI [m^2/m^2], PWC= plant water content [kg/ha]

With the following terms: the vegetation index WDVI, σ^0 the backscattering coefficient [linear], lia the local incidence angle [radian], nd the precipitation sum [mm], the coefficients $WDVI_\infty$, α, A, B, C and E. In order to contribute to the different modelling performances of the single sensor analysis, weighting factors were used. W_{CLAIR} (Fig. 4.6) stands for the coefficients of determination, when modelling with the CLAIR model, either before or after heading/flowering and W_{ML} for the multiple regression coefficients of determination for the respective biophysical parameter (Fig. 4.14). The previous findings for each sensor and biophysical parameter are taken into account through the different weighting factors. The fitted coefficients for this modelling approach are listed in Tab. A.11 within the chapter **ATTACHMENT**(A).
Due to the fact that the ASAR time-series was not dense enough, only three combined data pairs (4.10) could be matched, covering six fields before heading and only four fields after heading. Therefore, the analysis could not be done for the ASAR sensor, due to statistical reasons.

In Fig. 4.23 the retrieval results with the 1:1 line, the R^2 and the rmse for each biophysical parameter are displayed. Generally, when comparing the findings with the combination of NDVI and the backscattering coefficients (Fig. 4.22), they have improved significantly for all retrieved parameters. Also when collating them with the Fig. 4.14, showing the multiple regression results for the ERS-2 analysis they have improved. Especially, the results for the LAI and the dry biomass feature a significant improvement. When comparing them with the multispectral findings, using the CLAIR model there is no clear trend of improvement. For the fresh biomass during early phenological stages even though the R^2 has increased slightly, the trend of underestimation has also increased rapidly. During late phenological stages there is no clear trend of the scattering phenomena, but the R^2 has decreased in comparison to using the CLAIR model alone. For this biophysical parameter the shown combined approach does not really improve the results and the multispectral retrieval alone outperforms it. For the dry biomass in contrast, this approach features a high improvement in comparison to the CLAIR model. Especially, for late phenological stages the scattering around the 1:1 line has decreased and the trend of underestimation can not be observed any more. Although the weighting of the ERS-2 information was rather low with just 0.28 it improved the findings. For the LAI retrieval during early

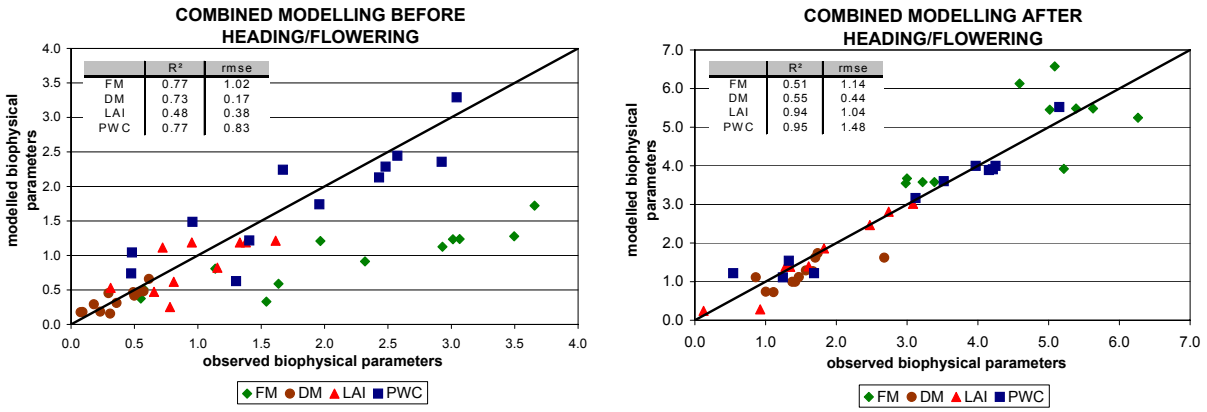

Figure 4.23: Combined biophysical parameter modelling using the CLAIR model and multiple regression, considering the WDVI (intercalibrated) and the ERS-2 backscatter. FM=fresh biomass [kg/m^2], DM= dry biomass [kg/m^2], LAI [m^2/m^2], PWC= plant water content [kg/m^2]

phenological stages, even though the scattering is not that high, the R^2 has decreased. During late phenological stages the findings are very good featuring a high R^2 and hardly any scattering. The actually little weighted (0.33) ERS-2 information has improved the performance. The modelling of the parameter plant water content has also improved for late phenological stages. Overall, especially during late phenological stages the ERS-2 information improved the modelling results in comparison to only using the CLAIR model (exception fresh biomass).

4.3.3 Combining the CLAIR and Water Cloud model

Finally the combination of the CLAIR and Water Cloud Model is analysed for all biophysical parameters. Again only the performances using the ERS-2 sensor could be assessed, due to the small amount of appropriate ASAR scenes. The fitted coefficients are in Tab. A.12 in the chapter **ATTACHMENT** (A). Biophysical parameter modelling was done by the term:

$$b_p = W_{CLAIR} \times \left[-\frac{1}{\alpha} ln \left(1 - \frac{WDVI}{WDVI_\infty} \right) \right] + W_{WCM} \times \left[\frac{(\sigma^0 - C \times \theta - D \times nd) \times cos(\theta)}{-2 \times B} \right] \tag{4.12}$$

with the vegetation index WDVI, σ^0 the backscattering coefficient [dB], θ the local incidence angle [radian], nd the precipitation sum [mm], the coefficients $WDVI_\infty$, α, B, C and D. For contributing to the performances of the different sensor systems weighting factors were used. W_{CLAIR} (Fig. 4.6) stands for the R^2, when modelling with the CLAIR model, either before or after heading/flowering and W_{WCM} for the R^2 of the performances from the Water Cloud Model, for the respective biophysical parameter retrieval (Fig. 4.18). The coefficients of determination are retrieved from the previous single sensor analysis. This way the single modelling performances for each respective parameter and sensor are taken into account. In Fig. 4.24 the respective results for each biophysical parameter are displayed. Generally, the results for modelling during early phenological stages are sounder than for the later stages. When, comparing the findings with Fig. 4.23 there is a decrease in the performance, except for the parameter fresh biomass there is an increase in modelling accuracy, especially concerning the scattering behaviour. But still the CLAIR model outperformed the combined approach for fresh biomass.

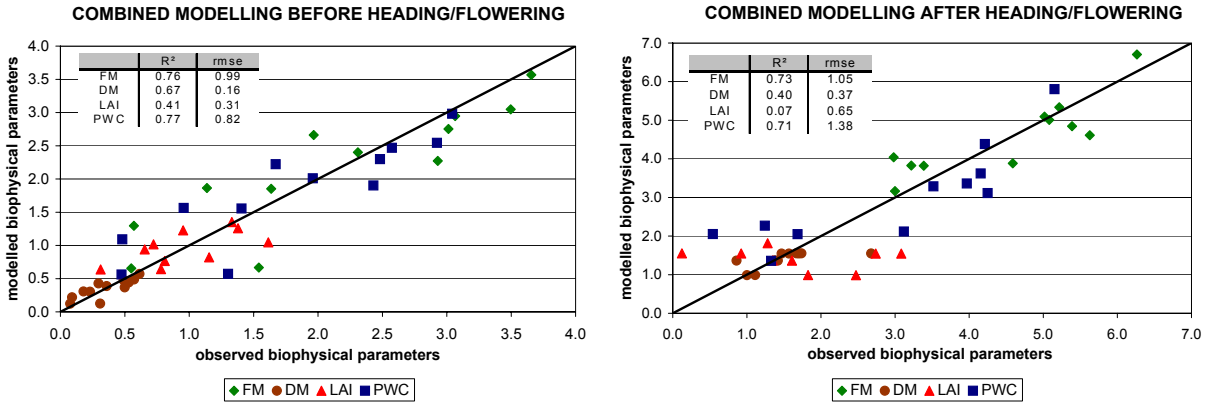

Figure 4.24: Combined biophysical parameter modelling using the CLAIR model and the Water Cloud model, considering the WDVI (intercalibrated) and the ERS-2 backscatter. FM=fresh biomass [kg/m^2], DM= dry biomass [kg/m^2], LAI [m^2/m^2], PWC= plant water content [kg/m^2]

4.3.4 Conclusions

Again within this section the overall question was:

Is it possible to derive the values of important crop variables from various remote sensing data?

here assessed in further detail:

- the potential of combining multispectral and SAR data for the derivation of biophysical parameters

- the inversion of the used models for regional biophysical parameter extraction

The analysed method of combining the multi-sensoral information by linear regression did not bring satisfactory results and was outperformed by single sensor use, only taking into account the multispectral information. Only for the fresh biomass modelling, using the NDVI and the ASAR ratio, a marginal improvement in regard to single multispectral sensor modelling was assessed. For the combined use of the CLAIR and Water Cloud Model no sound results were achieved. The combination by using the CLAIR model and multiple regression featured slight improvements, in contrast of the single multispectral sensor use. Especially, during late phenological stages the assessed VV information improved the modelling results, in comparison to only using the CLAIR model. Due to the use of rather simple input parameters it was also possible to enable regional parameter modelling.

4.4 Summary of the previous findings

When comparing all the various modelling approaches, using the different sensor systems for the four analysed biophysical parameters, there appears to be no generalized answer concerning the best sensor (polarization) to use. For the fresh biomass multispectral data using the CLAIR model and the ASAR multiple regression analysis, based on the HV polarization performed valid. The multispectral sensors outperform the ASAR HV polarization, in regard to low biomass and rather high biomass value (above 5kg/m^2) modelling. The ASAR HV polarization performs sounder in the biomass range between 4 - 5kg/m^2. Using the ERS-2 VV polarization functions satisfactory and modelling can be done over the whole biomass range. Concerning the parameter dry biomass modelling by the use of the VV polarization is not valid. The ASAR ratio exceeds the results using the CLAIR model before and after heading/flowering. Even so the scattering in the low biomass range for the multispectral sensor modelling seems to be a bit less. There is a clear trend of underestimation for modelling the high biomass range, where the SAR ratio outranges the multispectral data. For the plant water content retrieval, the multispectral data clearly outperform the SAR data for all considered times. But also modelling by the use of the VV polarization is appropriate. For the LAI retrieval the multispectral data deliver the best results, over the whole value range, with a slight underestimation in the high LAI range. The HV polarization is valid for LAI retrieval, however the scattering around the 1:1 line is dominanter than for the multispectral data. By just using the Water Cloud Model alone the retrieval results are not as sound.
The first assessed fusion method, by simple linear regression performed more valid, than the single multispectral modelling in the case of ASAR ratio combined with the NDVI, for the fresh biomass retrieval. Concerning all other biophysical parameters and combinations, the single multispectral modelling approach performed sounder. Regarding the other fusion approaches, the presented retrieval model combining the CLAIR model (multispectral data) and multiple linear regressions (SAR data) exceeded the synergetic modelling using the CLAIR and the Water Cloud model. The used weighting factors, coming from the previous single sensor analysis, improved the results. Especially, when the mutispectral information is weighted slightly higher than the SAR information. The suggest combined modelling approach improved the retrieval results in contrast to single sensor modelling, especially during late phenological stages. Large improvements were assessed for the dry biomass modelling especially, during early phenological periods, but also for stages after heading/flowering. The findings might even be improved by tuning the weighting factors, but this might be critical because there is no scientific reason. It would more or less be a try and error fitting of the coefficients. By the use of the different ASAR polarizations combined with the multispectral data the results of the suggested fusion approaches also might improve.

5 CERES-WHEAT CROP GROWTH MODELLING AND FINAL YIELD ESTIMATION

Within the following section the overall assessed question was:

Is it possible to integrate important crop variables gained from multi-source data into a dynamic crop growth model in order to increase the final yield estimation accuracy?

Challenges hereby were:

- to assimilate biophysical parameter maps retrieved from different remote sensing data into a crop growth model in order to improve the accuracy of wheat yield predictions (section 5.4)
- to analyse the effect of assimilating separately or synergistically SAR and multispectral time-series information (section 5.4)
- to analyse the effect of acquisition time and frequency of the assimilated multispectral data (section 5.5)
- to assess whether time gaps in the multispectral data series can be filled by SAR information (section 5.6)
- to transfer the findings of the point-based crop growth model for regional final yield prediction (section 5.8)
- to analyse the performance of the suggested approach in comparison to a simple direct yield estimation method (section 5.9)

The retrieved remotely sensed information will finally be assimilated into the CERES-Wheat model for crop growth modelling and final yield estimation. Therefore, first additional information will be prepared and the genetic coefficients listed in Tab. 2.3 will be calibrated. The used biophysical parameters dry biomass and final yield are converted into kg/ha (1kg/ha equals 10,000kg/m^2), due to comparability reasons with the CERES-Wheat outputs. In order to use the suggested assimilation method of input parameter re-initialization, a sensitivity analysis of the initial model parameters will be performed. The CERES-Wheat modelling results, when assimilating the different remote sensing information, will be analysed. The advantage of the suggested assimilation is that the retrieval errors from the remotely sensed parameters are not

that drastically forwarded into the model, as by, e.g., using the updating or forcing method. This is especially important, when assimilating the parameters gained from SAR data, as these findings were not that sound. Then a sensitivity analysis of frequency and acquisition time of the multispectral data will be assessed, due to the fact that multispectral data acquisitions are often hampered by cloud coverage. The question hereby is, if SAR information can overcome a lack of multispectral data. Finally, the data set performing best in the assimilation process in regard to the final yield, will be used for crop yield estimation within the whole research region. These results will be then compared to an approach of direct multi-sensoral yield estimation.

5.1 Additional information

Firstly, the additionally needed information had to be prepared, which were multi-temporal inputs, e.g., weather information and spatially varying inputs, e.g., soil information. Also general assumptions about the management practices, e.g., row spacing were made on the basis of the in-situ information observed in the research years. Concerning the set management practices the row spacing was set to 12cm and the planting depth to 5.5cm. The daily weather data, measured at the climate station in Klein-Altendorf from 2004 until 2006 were transformed into daily temperature minimum and maximum [°C], daily precipitation [mm] and daily values of incoming solar radiation [MJ/m^2-day]. The soils were grouped into four major classes, which were: Luvisols, Cambisols, Gleysols and Stagnic Gleysols. For each field the soil type with the largest surface fraction, of that respective field, was taken. Slope aspect, elevation and obstruction to sun information were extracted from a digital elevation model.

5.2 Calibration of genetic coefficients

An important step when using the CERES-Wheat model, is the model calibration, i.e. the estimation of the genetic characteristics (Tab. 2.3) reported in the genotype input file, in order to obtain accurate predictions. The calibration of the CERES-Wheat model was performed by using the experimental data collected over the two winter wheat fields monitored in 2005.
For the determination of the genotype parameters of the wheat varieties typically sown in the region of Meckenheim, the modelled temporal course of the dry biomass divided into stems, leaves and ears, the total above ground biomass, the LAI, the final predictions of total above biomass, the final yield and the flowering dates were compared with the in-situ data. For the assessment the simulation error, e.g., the mean absolute difference between the estimated and the observed data were considered. Then the genetic coefficients (Tab. 2.3) were varied, in order to minimize the differences between the simulated and measured values within acceptable error limits. For this procedure the **GENCALC** program utility (Hunt et al., 1993) was used, varying first the coefficients regarding the phenology (P1V, P2D, P5 and PHINT) and in a second step the production related parameters (G1, G2 and G3). Fig. 5.1 shows exemplarily for 2005 the finally modelled and observed total dry biomass development.
After calibration the flowering dates were simulated exactly by +/- 0 days. The measured R^2 and rmse for the dry biomass were for field 1: 0.952 and 1662kg/ha and for field 2: 0.934 and 1830kg/ha (Heinzel et al., 2007). Final grain yield differences between modelled and observed

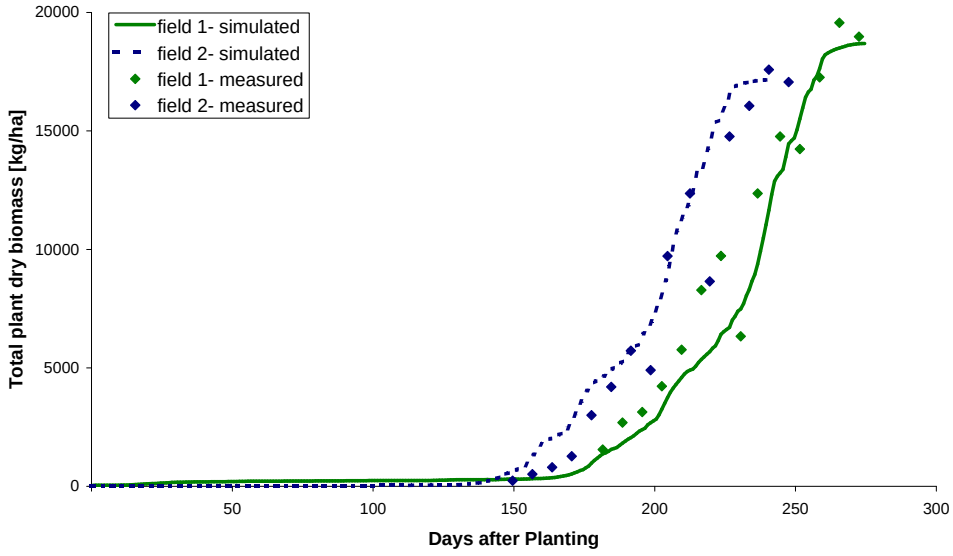

Figure 5.1: CERES-Wheat calibrated simulation results and measured dry biomass [kg/ha], exemplarily shown for the two winter wheat fields observed in 2005.

were 985 and 694kg/ha, having an error of 12% and 7%. Compaired to Dente et al. (2007), who achieved a yield prediction error of 24% for the calibration process of winter wheat in Matera (Italy), the results are very sound.

To validate the prediction capabilities of the model, a different independent set of experimental data, consisting of data collected from four winter wheat fields in 2006, were compared with the simulated parameters. The measured R^2 for the four fields was 0.91 (yield) and the final grain yield difference was 1080kg/ha, which was an error of 13%. This indicates a sound quality of the model estimates, also being able to account for wheat variety and different climatic conditions.

5.3 Sensitivity analysis of the initial model parameters

Even though the genotype parameters are identified during the calibration procedure, the CERES-Wheat model still requires a large amount of input parameters. Among them, climatic data, which were less critical to obtain, due to the meteorological station. But the model also needs management information, which can be largely variable. Under these circumstances, a crucial point is to determine whether or not there is a relatively small subset of model inputs, with especially high impact on the modelling output. This subset would influence more than others, the temporal behaviour of the state variables of interest. Consequently, a sensitivity study was carried out in order to identify initial conditions, having the highest influence on the final results, which was done exemplarily for the monitored fields in 2005. In regard to this different management characteristics (sowing and nitrogen fertilization date, planting density) were analysed. The analysis did not include climatic inputs, which were considered known, as well as

plant density and row spacing, which were almost constant in the area and soil characteristics, which were given by the geological survey North Rhine Westphalia (Dworschak et al., 2001) and in further detail for the research farm by Lehmann & Pätzold (1996).

In a first step the parameter planting date was varied between the 10/15/2004 and 01/01/2005, with constant other parameters. The sowing date is an important management feature, as it establishes the period of the season, when wheat growth occurs and then, depending on the temperature of that period, the length of the phenological stages. For the second sensitivity study the parameter plant population at seeding was varied between 300 to 450 plants/m^2, keeping constant other input parameters. In a third study the dates of the fertilizer applications were varied in two weeks steps for each of the three respective treatments. Starting from 1. March, 1. April and 1. May for the respective three applications (early version) until 1. April, 1. May and 1. June (late version) for the three applications, but always applying a constant amount of nitrogen fertilizer, 180kg/ha in total. In Tab. 5.1 the variations within the final yield for the three sensitivity studies are shown. In relation to the actual final yield (9850kg/ha) the variations within the different planting dates are from +0.5% to -12%, for the different plant populations 0.5% to 3.9% and for the different fertilizer application dates from -5% to +12%. Overall, the variation of the different resonable planting dates featured a high variation from 12.5% in the final yield and the different fertilization dates from nearly 17%, thus these two seem to have the highest impact on the modelling performance (Heinzel et al., 2007).

Table 5.1: Final yield results [kg/ha] for varying planting dates ($dates$), planting density ($density$) and fertilizer application dates ($application$).

dates		density		application	
10/15/2004	9897	300	10236	early version	11536
11/01/2004	9456	400	9897	mid version	9897
11/15/2004	9099	450	9780	late version	9434
12/01/2004	8755				
12/15/2004	8853				
01/01/2005	9411				

The other important unknown parameters were set to a constant value, e.g., the plant density to 300 plants/m^2 with a row spacing of 12.5cm. Tillage was also considered to be constant for the whole research region and exemplarily taken from the research fields. Tillage was considered during autumn, first with a row crop cultivator (tillage depth 15cm) and than with a row crop planter (tillage depth 2cm). Chemical applications were given automatically when necessary. Organic amendments were not considered, as well as irrigation, which is not necessary in the research region.

5.4 CERES-Wheat modelling by assimilating remotely sensed information

The actual crop growth simulation was done for the research year 2006, thus having a totally independent data set, where no CERES-model calibration was conducted. Firstly, modelling

was only done for the four wheat fields monitored on the research farm, in a second step for the best data assimilation results the findings were transferred to the whole research region. The assimilation method adopted (section 2.3.4) performs a re-initialization of the CERES-Wheat model. It consists of a procedure that tunes the model initial conditions, until the temporal behaviour of the model state variables, e.g., LAI or dry biomass reach the best agreement with the multi-temporal remotely sensed information. It is assumed that the CERES model and the retrieval algorithm estimating the state variable values are not biased. In order to search for the optimal configuration of the model the mean simulations and observations, the mean ratio, the standard derivations of both, the R^2, the mean difference, mean absolute difference, the rmse and the flowering dates were taken into consideration. Additionally, also the original retrieval performances from the remote sensing data were considered. In regard to their temporal behaviour initial conditions were set. There was a constraint on the variability ranges of these parameters, in order to avoid unrealistic conditions and to take into account the site characteristics, considering all the information, concerning the respective parameters collected in the two years of campaigns, into account. The error on the initial information was set equal to half of the variability range of these parameters. The initial sowing dates ($sowing$) were fixed at the beginning of November with a standard variation of approximately one month (the observed variability was from beginning of October until first of December). The fertilization dates ($fertilization$) were varied within the in the sensitivity study used range (early, mid and late version). In general there were 16 simulation runs performed. In the first five variable planting dates in combination with the early fertilization option were varied, for the next five the variable planting dates were combined with the mid fertilization dates, in the next five the planting dates were combined with the late fertilization dates. For the 16 simulation run the in-situ information, gained from the local farmer were used. After determining $sowing$ and $fertilization$, the CERES-Wheat model was re-initialized with the respective optimum set of input parameters.
The described assimilation method was first assessed for the research farm and afterwards extracted to the whole region considering the remote sensing information set performing best in regard to final yield estimation. The used biophysical parameter maps were dry biomass and LAI maps retrieved from the respective remote sensing data. Generally, first the dry biomass maps were considered and in a second step the retrieved LAI maps, but also considering the respective retrieval performance from the remote sensing data. In case of large differences between the retrieval performances the parameter with the better results was weighted higher. For the ERS-2 data this can be quite problematic due to the poor retrieval results featuring only a $R^2 = 0.33$ for the LAI and a $R^2 = 0.28$ for the dry biomass. The following databases were tested for the model re-initialization:

- automatic sowing and fertilizer date setting, which will be named *automatic yield*
- the biophysical parameter maps retrieved from multi-temporal multispectral data using the CLAIR model (before and after heading), named *CLAIR yield*
- the biophysical parameter maps retrieved from the ERS-2 time-series using multiple regressions, named *ERS-2 yield*
- biophysical parameter maps retrieved from the ASAR data using the respective best polarization and the multiple regression approach, named *ASAR yield*

- biophysical parameter maps retrieved by combining the multispectral CLAIR model with the multiple regression for the ERS-2 data, named *Combined yield*

- time-series of biophysical parameter maps retrieved from ERS-2 and ASAR data, named *SAR time-series yield*

- time-series of biophysical parameter maps retrieved from all available SAR and multispectral data, named *Multispectral and SAR time-series yield*

5.4.1 Automatic sowing date and fertilizer date setting (*automatic yield*)

In a first step CERES-wheat yield modelling was done without considering any further information. Sowing dates and fertilizer dates were set automatically, by the DSSAT software. Sowing was performed on the four fields within the period of 09/01/2005 and 12/01/2005, when the soil temperature was around 10C and the soil water content was reasonable for sowing. Fertilization was done when needed, with a threshold of N stress factor at 50.
The automatic sowing for the four analysed fields was 10/01/2007 with 180kg/ha fertilizer applied and final yields of 7273kg/ha. There were only slight simulation differences between the four fields due to very similar soils and the same weather data. The rms error between *automatic yield* and the measured yield is rather high with 629kg/ha and a low R^2 of only 0.07. An overestimation of 6.6% of the final yield occurred averaged over the four fields. This proves the assumption that it is rather critically to model final yield without a sufficient amount of additional information.

5.4.2 Multispectral information (*CLAIR yield*)

The LAI and dry matter maps retrieved by the different multispectral images from 2006 (Tab. 3.4) provided a temporal series from stem elongation to ripening during the wheat growth period, to estimate for each field the optimum input parameters and re-initialize the CERES-Wheat model. The LAI estimates errors from the remote sensing data were 0.42 and $0.43 m^2/m^2$ (rmse) ($R^2 = 0.85/0.65$), for the dry matter estimates errors were 1600 and 5400kg/ha (rmse) ($R^2 = 0.58/0.54$). These errors will also influence the model performance and have to be taken into account when adjusting the input parameters to the remotely sensed information. Due to the slightly sounder findings for the LAI retrieval this parameter was weighted higher.
In Fig. 5.2 exemplarily the modelled LAI and dry biomass (wheat field 1) using variable planting dates and fixed fertilization dates (mid version) are shown. Additionally, the achieving using the in-situ data together with the respective remotely sensed variables are listed. By only regarding the different curves it is hard to conclude, which setting would be the optimal solution. In a second step the statistics of the models were analysed taking especially into account the different rms errors (Tab. 5.2). What becomes apparent is that the LAI rms varies between 1.1 and 0.6 and the dry biomass rms between 4373 and 6568kg/ha, however for both parameters the lowest values were obtained for simulation 11. The model for wheat field 1 is re-initialized using the planting date 10/01/2005 and the late fertilization option. This means a bias to the

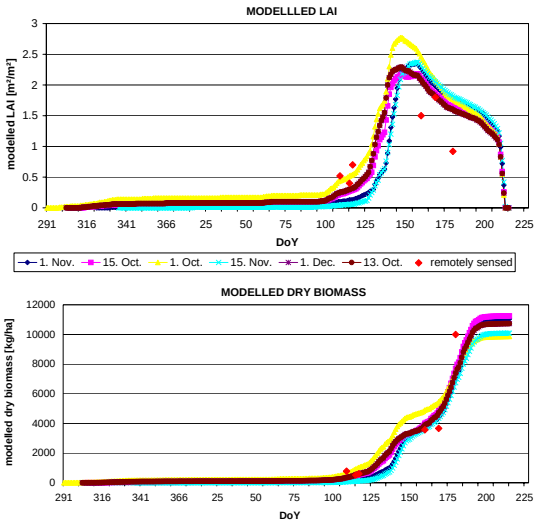

Figure 5.2: Modelled LAI $[m^2/m^2]$ and DM [kg/ha] with variable planting dates, plotted against the days of the year (DoY).

Dates for fertilization were set (mid March, mid April and end of May) for this modelling case. Results for the 10/13/2005 were achieved by using the ground truth information for the local farmer.

in-situ data of 12 days for the planting (to early) and fertilization is delayed for all three applications around two weeks. However, the final yield is only overestimated by 15kg/ha, which is a splendid result. Overall, fertilization dates were after the assimilation process in rather good

Table 5.2: Rms errors for LAI and dry biomass [kg/ha](wheat field 1) and the final yield [kg/ha] using the 16 simulation runs and the remotely retrieved parameters.

run	LAI rms	DM rms	yield	run	LAI rms	DM rms	yield
1	1.1	4385	7415	9	0.8	6097	9036
2	1.1	4843	7703	10	0.9	6568	9439
3	0.9	5766	8650	11	0.6	4373	7455
4	0.9	5721	8495	12	0.9	4732	7682
5	0.9	6198	8928	13	0.7	5466	8429
6	0.9	4919	8349	14	0.7	5411	8211
7	0.8	5239	8280	15	0.7	5748	8470
8	0.8	6169	9240	16	0.9	4943	7896

agreement with the in-situ data. Only for field 1 the late version was chosen. But the planing dates varied by nearly one month, i.e. the estimated sowing dates were delayed with respect to the observed ones.
Generally, when using the suggest assimilation strategy for the four monitored fields there was an average yield of 7716kg/ha with a rms error between the simulation and the real average of 562kg/ha. This is a slight yield underestimation of 0.9%, when using six multispectral data sets, covering the wheat growing season from stem elongation over flowering to ripening. Concerning the variabilty of the bias for the respective fields it is rather high. For the first three fields there was a slight overestimation (between 15 to 850kg/ha). Regarding the last field, featuring

a rather high in-situ yield there was a big understimation obvious (1548kg/ha). In this case the model can not simulate the observed yield variations. This might be due to the biases within the remotely sensed parameters, because especially for high amounts of dry biomass and LAI the retrieval performance was not that sound, featuring a slight saturation of the remote sensing signal. Or the calibrated genetic coefficients were not appropriate for high final yield modelling. A deeper analysis about yield variation estimates goes beyond the limit of this work, because it would require a higher number of monitored fields featuring a larger dispersion between the measured yields.

5.4.3 Information retrieved from ERS-2 data (*ERS-2 yield*)

The ERS-2 data for the retrieval of the LAI and dry biomass are listed in Tab. 3.6. The six data sets cover the wheat growth period from stem elongation over stem elongation and heading to the ripening, in contrast to the multispectral data they cover a longer time period. Generally, the errors of the LAI and dry biomass estimation were rather high. For the dry biomass the error amounts to 5600kg/ha ($R^2 = 0.28$) and for the LAI $0.62 m^2/m^2$ ($R^2 = 0.33$), the relationship between modelled and observed parameter were slightly higher for the LAI, therefore it will be decisive.
When assimilating these data into the model it was found that compared with the in-situ data the fertilization dates were in good agreement. The planting dates varied especially for two fields, there was a bias of nearly one month, once being delayed and once being to early. For the other two fields there was only a bias of +/- 2 days.
The estimated yield map for the four fields featured an average yield of 7928kg/ha with a rms error of 612kg/ha and a yield overestimation of 1.8%. Concerning the variabilty of the bias for the individual fields it is rather high. While the first three fields are overestimated between 387 and 816kg/ha, again the field with the high measured yield features a high underestimation of around 1082kg/ha. Actually the trends of over and underestimation from the individual fields sum up to a valid average final yield estimation. In regard to the remotely sensed parameters as for the multispectral data the bias in the high value ranges was bigger featuring an underestimation of the respective parameters. Generally, the findings for assimilating ERS-2 data into the model are sound, especially when considering the poor retrieval performance from the remote sensing data.

5.4.4 Assimilation of information from ASAR data (*ASAR yield*)

In a third step of the analysis the biophysical parameter maps retrieved from the ASAR data were assimilated into the model. The five ASAR data only cover a rather short crop growth period with some acquisitions at stem elongation and some at the important stage of fruit development. For the dry biomass the error amounts to 6700kg/ha ($R^2 = 0.82$) and for the LAI $0.6 m^2/m^2$ ($R^2 = 0.6$). The relationship between modelled and observed parameter were slightly better for the dry biomass, therefore it will be decisive. Again for this simulation setting for all four fields the mid version of fertilization dates was picked, but the planing dates varied again. Especially, for field two, which was sown on 11/14/2005 there was a bias of nearly one month, while for the other three the bias was only around 1-2 weeks.

The averaged estimated final yield was 8013kg/ha featuring a rms error of 621kg/ha, which is an actual overestimation of the final yield by 2.8%. The variablity of the results is a bit different than for the previous findings. Now there was a yield underestimation for field 1 (73kg/ha), while for the other three fields there was a trend of overestimation obvious. For field 2 the bias was 436kg/ha, which is slightly higher than with the assimilated ERS-2 information. Now there was a high overestimation of the final yield obviuosly for field 3 (1236kg/ha) and an underestimation of 699kg/ha for field 4. This trend can be explained by the retrieval performances, when using the ASAR data (Fig. 4.15). Especially, for the dry biomass extraction there was no clear trend of over or underestimation obvious. The LAI retrieval featured a scattering in the mid value range. The problems of a high value underestimation for high values, described previously, is not obvious. For the field 3 probably the high overestimation is due to the higher weighting of the dry matter findings. While for all other fields LAI and dry matter findings are normally in good agreement with each other, for this field there was a high bias between them.

5.4.5 Assimilation of synergetically retrieved biophysical parameter maps (*Combined yield*)

In a next analysis step biophysical parameters retrieved synergetically from multispectral and ERS-2 data (Tab. 4.9) were assimilated into the crop growth model. The time-series for 2006 only includes four data pairs from May until mid July, thus being not very dense. For the dry biomass the error amounts to 1700kg/ha ($R^2 = 0.73$) for early phenological stages and 4400kg/ha ($R^2 = 0.55$) for late stages. For the LAI $0.38 m^2/m^2$ ($R^2 = 0.48$) and $1.04 m^2/m^2$ ($R^2 = 0.94$). The relationship between modelled and observed parameter were slightly better for the dry biomass, thus it will be decisive. Considering the chosen initial settings for all four wheat fields the mid fertilization dates and the planting date 10/01/2005 was taken. This means that the planting date picked was to early for all fields. The average final yield was 7367kg/ha with a rms error of 585kg/ha, which are around 5.4% of yield underestimation. Again, especially for wheat field 4, having the highest real final yield the bias is rather high with an underestimation of 1548kg/ha. For field 3, with the lowest measured final yield, there was again an overestimation obvious of 388kg/ha.

5.4.6 Assimilation of the information from the SAR time-series (*SAR time-series yield*)

The big difference to the previously described assimilation procedures is that the temporal sequence of the SAR time series is much denser now covering 11 acquisitions (Tab. 3.6 and Tab. 3.7). The retrieval biases from both SAR data sets now have to be taken into account. Due to the lower retrieval bias of the ASAR data, these will be weighted slightly higher as well as the dry matter, for which the retrieval performance using ASAR data was more stable.
Considering the fertilization date again the mid version was chosen. Planing dates picked were correct for field 1, two weeks too early for field 2 and around 40 days to late for the fields 3 and 4. The final averaged yield was 8044kg/ha with a rms error 580kg/ha, hence an overestimation of 3.3%. As for the other data assimilations there was a yield overestimation for the first 3 fields

and a rather large yield underestimation (1081kg/ha) for the 4 wheat field. Compared, to using the ERS-2 or ASAR data individually, the bias slightly increased.

5.4.7 Information of the multispectral and SAR time-series (*Multispectral and SAR time-series yield*)

In a final data assimilation examination all available data sets from 2006 were taken into consideration for the re-initialization of the CERES-Wheat model. By using this assimilation strategy the temporal sequence of the biophysical parameter is the densest covering the whole growing season with 17 single acquisitions. However, also the retrieval errors from the different remote sensing data have to be considered now and will probably be resumed in the re-initialization process. Due, to the fact that the assimilation performance using the multispectral information outperformed all other tested runs the multispectral information (especially the LAI) will have the highest weight. As already previously, for the fertilization dates the mid version was chosen. Concerning the planting dates they were set in all cases around one month later than the in-situ observations. There was a slight yield overestimation of around 0.56%, with an average yield of 7833kg/ha (rmse=541kg/ha). The trends for the individual monitored fields were the same as for the previously described assimilation strategies.

5.5 Sensitivity to frequency and acquisition time of the multispectral data

Due to the fact that multispectral acquisitions are often hampered by frequent cloud coverage, the assimilation performance was evaluated with different time gaps in the multispectral data series. In a first step single multispectral acquisition dates were left out analysing the impact when specific inputs from distinctive phenological periods were missing. The data series (Tab. 3.4) contains three scenes during stem elongation, one during flowering and two during the ripening process. It is expected that especially when leaving out the remote sensing information from dates with maximum LAI, should have an impact on the assimilation performance, as this are distinctive periods during crop growth. When using all multispectral data sets for the assimilation there is a slight yield underestimation of 0.9%. When neglecting the first acquisition, there is a final yield overestimation by 3.1%, for leaving out the second acquisition there is an underestimation of 0.8%, for neglecting the last acquisition from the stem extention period an overestimation of 0.5% occures. Leaving out the remotely sensed information during flowering an underestimation of 0.8%, and for the last two an overestimation of 0.5% and an understimation of 0.8% occure. It seems that, especially the first acquisition is rather important for an exact yield estimation. During other phenological stages a lack of one multispectral image actually improved the yield estimation in regard to using all data, regardless which acquisition date was missing. The rms error between the modelled and remotely sensed biophysical parameters also did not vary that much for these last five assimilation runs. Generally, it can be concluded, that actually a too dense data series declines the final crop yield estimation. This is probably due to the different retrieval errors, when estimating the biophysical parameters from the remote sensing data. They are also introduced into the assimilation process. The impact of the absence

of one out of six images is negligible, as long as one image during an early phenologcial stage is present. However, this is not really a conclusion to be generalized due to the fact that actually acquisitions were not that good distributed during the whole crop growth period. Another fact is the large impact of the first date, which should not be that high, due to the fact that there are also two more acquisitions during this phenological stage and because it is not acquired in the period of maximum LAI, which is one of the most important stages for the determination of the final yield.
In a next step the model performance was evaluated with data gaps from specific phenological stages (Tab. 5.3). Firstly leaving out all data during stem elongation, then during flowering and finally during the ripening stage, in order to actually assess the importance of data during specific phenological stages. When the acquisitions during stem elongation were missing, the effect of the data lack is more pronounced. The total final yield was underestimated by 5.4% (7367kg/ha). When neglecting the multispectral image during flowering the final yield underestimation was not as high with only 0.8% (7725kg/ha). Leaving out the last acquisitions during ripening only results in a final yield overestimation of 0.56% (7833kg/ha). This trend is also obvious in the LAI and dry matter rmse between the modelled and remotely sensed information. Especially, the LAI rmse was rather high when leaving out the images acquired during stem elongation. For the dry matter this trend is not that distinctive. The approach leaving out images during ripening produces the smallest LAI rms, featuring a rather good agreement.
This shows that especially early acquisitions from stem elongation have a higher impact on the final yield modelling results, than images after this stage. Especially, images during the ripening seem to hardly influence the final yield estimation. When leaving out these images final yield estimation bias was rather small. In regard to using all available data sets final yield estimation seems to actually improve when leaving out the information during flowering or ripening. Compared to the achievements using the *SAR time series yield* approach the final yield estimation performance only decreases when leaving out the data set during stem elongation. For all other time gaps the multispectral data assimilation still outperforms the SAR data series or ERS-2 or ASAR data assimilation method.

5.6 Filling multispectral acquisition time gaps, by considering SAR information

Finally, the sensitivity of the assimilation procedure to a lack of multispectral data in the different periods of the wheat growth season, considering additionally SAR information, was evaluated. Multispectral data acquisitions are often hampered by cloud coverage, thus it is of interest if SAR information are able to overcome these time gaps. The time-series of SAR and multispectral data was used, with simulated gaps in the multispectral time-series. Firstly neglecting all data during stem elongation, then during flowering and finally during the ripening process. The question hereby was, if multispectral time gaps, especially during stem elongation, are negligible when a large number of SAR data are available for obtaining a reasonable result.
In Tab. 5.4 the rms errors between the simulated and the remotely retrieved LAI, dry matter as well as the modelled final yield and the percental bias to the real observed yield are listed. In total 240 individual simulation runs were performed for the four monitored winter wheat fields varying the planting dates (5 different dates), the fertilization dates (early, mid and late version) and missing multispectral data from distinctive phenological stages (3 phenological

Table 5.3: Rms errors for LAI and dry biomass [kg/ha], final simulated yield [kg/ha] and yield bias in regard to in-situ measurements [%].

run	wheat field	LAI rms	DM rms	final yield	yield bias
run 1	1	0.6	4373	7455	0.2
	2	0.5	5283	7455	-4.6
	3	0.6	5605	7833	12.2
	4	0.6	6083	7833	-12.1
run 2	1	0.84	5457	7367	-0.98
	2	0.83	4660	7367	-5.79
	3	0.86	5417	7367	5.54
	4	0.83	5160	7367	-17.36
run 3	1	0.57	5345	7833	5.28
	2	0.49	4757	7833	0.16
	3	0.56	5363	7833	12.22
	4	0.57	6041	7403	-16.96
run 4	1	0.33	5415	7833	5.28
	2	0.50	6051	7833	0.16
	3	0.38	5669	7833	12.2
	4	0.36	6506	7833	-12.1

run 1= all multispectral data, run 2= only multispectral data from flowering and ripening stage, run 3= multispectral data from stem elongation and ripening stages, run 4= multispectral data from stem elongation and flowering stages

stages) used for the assimilation process, whereby in Tab. 5.4 only the simulation statistics of the chosen run are shown. Using the SAR and multispectral data series without multispectral information from stem elongation again featured the largest final yield bias with an overestimation 5.8% (average final yield: 8246kg/ha). The bias has even increased in regard to only using multispectral images from flowering and ripening. The SAR information was not able to fill this important time gap. The same can also be observed for leaving out the multispectral information during flowering stage. Then the bias increased to 1.8% (average final yield: 7928kg/ha), which was an overestimation 1% higher than only using the multispectral information during stem elongation and ripening. Leaving out the multispectral information during ripening in combination with the SAR information producess the same bias (0.56% overestimation) than modelling without additional SAR information and no further multispectral information about the ripening stage.

Concluding one can say that the SAR information is not appropriate for filling missing time gaps in the multispectral time-series. It more appears that the modelling performance even decreased in regard to using a multispectral data series containing time gaps of distinctive phenological periods.

Table 5.4: Rms errors for LAI and dry biomass [kg/ha], final simulated yield [kg/ha] and yield bias in regard to in-situ measurements [%].

run	wheat field	LAI rms	DM rms	final yield	yield bias
run 1	1	0.54	5263	7833	5.3
	2	0.51	5431	7833	0.2
	3	0.5	5388	7833	12.2
	4	0.8	4604	7833	-12.1
run 2	1	0.54	5009	8256	10.9
	2	0.51	4976	8256	5.6
	3	0.57	5144	8256	18.3
	4	0.51	4838	8216	-7.8
run 3	1	0.53	5180	7833	5.3
	2	0.47	5265	7833	0.2
	3	0.47	5260	7833	12.2
	4	0.54	4685	8126	-7.8
run 4	1	0.47	5180	7833	5.3
	2	0.47	5520	7833	0.2
	3	0.4	5386	7833	12.2
	4	0.39	5717	7833	12.1

run 1= all SAR and multispectral data, run 2= all SAR, multispectral data from stem elongation left out, run 3= all SAR, multispectral data from flowering left out, run 4= all SAR, multispectral data from ripening left out

5.7 Discussion and conclusions

In the previous section a method to assimilate remotely sensed dry matter and LAI maps into the CERES-Wheat crop growth model, in order to improve the accuracy of the final wheat yield estimates, was assessed. The assimilation strategy consisted of tuning a set of management characteristics, until the differences between the modelled LAI and dry matter curves and the remotely retrieved biophysical parameters were minimized. Then the CERES-Wheat model was re-initialized using the optimal input parameters. A preliminary sensitivity analysis of the CERES-Wheat model helped to identify the most suitable model inputs to be involved in the re-initialization. It showed that planting and fertilization dates mostly effected the LAI and the dry matter development and therefore the final yield.
The method was firstly applied and validated with the four monitored ground truth wheat fields (2006) in Klein-Altendorf.
After the assimilation of the six different data series and modelling using an automatic setting, yield maps with different accuracies were obtained (Fig. 5.3 and Tab. 5.5). When using an automatic model setting the achieved results were biased with a yield overestimation of 6.6% over the four monitored fields. The algorithm integrated into the DSSAT software for choosing the right settings did not performe very well in the case of Klein-Altendorf. Regardless, which other data series were assimilated into the CERES-Wheat model, the final yield estimation accuracy increased. The assimilation of the six LAI and dry matter maps, retrieved from the multispectral data, led to a wheat yield map featuring an estimated rmse of 562kg/ha and an average yield of 7716kg/ha. In contrast to the automatic yield modelling the estimated final yield bias decreased to an underestimation of only 0.9%. The results obtained for assimilating the ERS-2 data series, were in comparision to the findings from *CLAIR yield* slightly poorer with an overestimation of 1.8%. The results achieved for *ASAR yield* featured a higher bias with a final yield overestimation of 2.8%. This is probably due to the fact that there were only four acquisitions available, covering a rather short phenological period. The actual biophysical retrieval results from this sensor were very precise. The finding for the *Combined yield* approach were not as good as the achievments, when using the multispectral data or SAR data alone, with a yield underestimation of 5.4%. Also featuring the trend of high yield underestimation for high in-situ measured yield (e.g., field 4). Using a denser SAR time-series (*SAR time series yield*) did not improve the final yield estimation accuracy, in regard to only using the ERS-2 or ASAR time series. There was now a final yield overestimation of 3.3% in contrast to only 1.8%, when only using the ERS-2 data or 2.8% for the ASAR data. The final yield estimation using the dense SAR time-series, led to a decreasing modelling performance. Overall, the best results could be obtained when using a very dense data series by combining all 17 available satellite data. The bias was then only 0.56% overestimation. This is a bit contradictornary to the findings from the multispectral time gap analysis. There the results actually improved, when leaving out data from specific phenological stages, as long as information from stem elongation was available. The sound result might be due to the fact that generally for the multispectral data series assimilation, there was a final yield underestimation, while for the SAR data series an overestimation trend could be observed.
However, except for the retrieval results from *ASAR yield* there were problems when estimating rather high final yields as for wheat field 4. They all featured then a clear trend of a high final yield underestimation. A deeper analysis about yield variations would actually be necessary, which goes beyond the limit of this work. It would require a higher number of monitored fields featuring a greater final yield variety. The calibrated genetic coefficients should actually allow

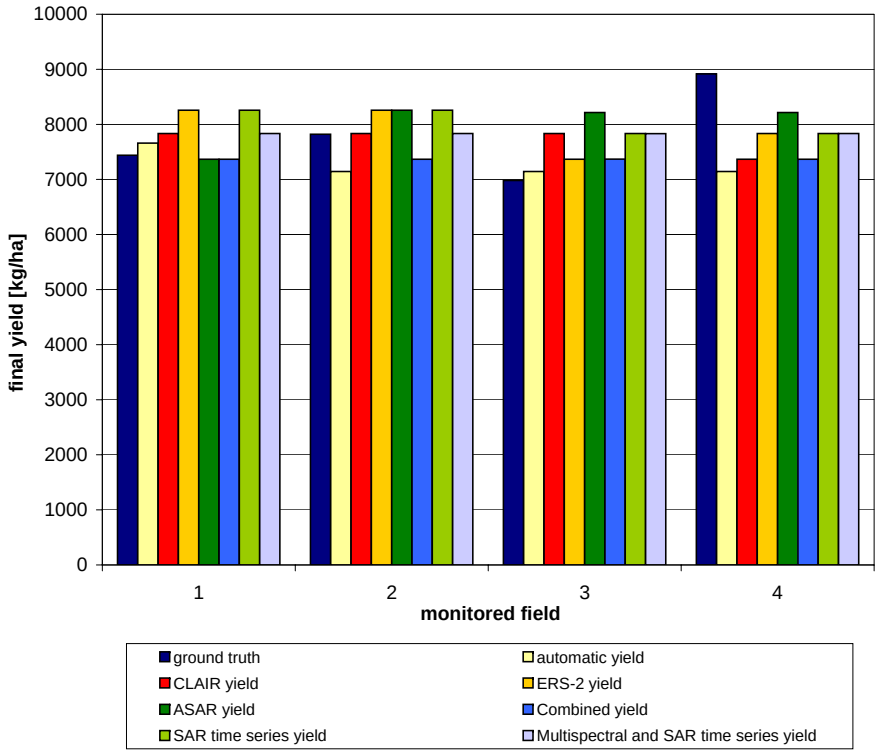

Figure 5.3: Final wheat yield estimates (fields 1-4), using the different data series for the assimilation procedures.

a higher yield estimates, which was done for field 4 in the *ASAR yield* assimilation procedure. The sensitivity of the assimilation process to acquisition time and frequency of the multispec-

Table 5.5: Statistics of the yield maps obtained after the assimilation of remote sensing data into the model.

assimilated data	no. of data	mean [kg/ha]	rmse [kg/ha]	bias [%]
ground truth		7788		
automatic yield		7273	629	6.6
CLAIR yield	6	7716	577	-0.9
ERS-2 yield	6	7929	612	1.8
ASAR yield	5	8013	621	2.8
Combined yield	4	7367	585	-5.4
SAR time-series yield	11	8044	580	3.3
Multispectral and SAR time-series yield	17	7833	541	0.56

tral data was also investigated. Multispectral data acquisitions are often hampered, due to cloud coverage. It is interesting to know how many multispectral images are needed and at which phenological stages, in order to improve the yield estimation accuracy in regard to just using SAR data. Results indicated that the lack of multispectral data in the assimilation process does not significantly effect the accuracy of the predicted yield maps, provided that the remaining assimilated maps cover the most important early wheat stages. The analysis proposed that the multispectral information from flowering or ripening stages only slightly influence the prediction performance, when information from stem elongation was present. It even appeared that a

too dense multispectral time-series had a negative effect on the final yield prediction accuracy. Results were preciser when not considering the full multispectral time-series, as long as the first LAI and dry matter maps were regarded from beginning of May. When trying to fill the gaps within the multispectral time-series by using SAR information, the precision for the first two approaches (no multispectral information from stem elongation or flowering) even decreased the final yield modelling performance. When leaving out multispectral information during ripening the result was comparable to the achievement without considering SAR information. The SAR information was not able to fill multispectral time gaps during important phenological stages.
Overall, the best yield prediction results were achieved, when using all available remote sensing data (SAR and multispectral data). The same final yield biases were also achieved, when using only multispectral information, without considering data from the phenological stage ripening. Also the same final bias was found when regarding all available SAR information together with the multispectral data, without the data from the phenological stage ripening.

One can conclude: Multispectral information from the phenologcial stage ripening actually had a negative impact on the final yield estimation accuracy. Information from this phenological stage led to a higher final yield estimation bias. When neglecting the multispectral information from this phenological stage, it made no difference if the SAR information was additionally taken into account. By just considering the multispectral information the same results were found as when also regarding the SAR information.

This is contradictory to the findings from Dente et al. (2007). They figured out that a lack of LAI maps from multispectral data during critical time periods can be overcome by additionally using SAR information, providing the same results as when using the whole multispectral data series. Within this study it more seemed that the additional SAR information actually declined the modelling performance, when multispectral information from stem elongation or flowering were missing. For neglecting multispectral information from the phenological stage ripening, the SAR information had actually no impact on the modelling result.
In conclusion the assimilation method described, is a promising technique to apply crop growth models, such as CERES-Wheat, when no further information is available and the management decisions are set automatically instead. The main critical aspect related to this is that the employed crop growth model needs a calibration of the genetic coefficients (2.3) for the assessed wheat variety. This implies that the illustrated approach might only be appropriate for agricultural areas relatively homogeneous in terms of wheat crop variety. This will be evaluated in a next step, applying the findings of the best assimilation approach, containing the least amount of information (multispectral data series, without phenological stage ripening) and an automatic setting approach to the whole region of Meckenheim. The results will then be compared to a direct yield estimation method.
Another more general aspect, which has to be kept in mind, is that it can be problematic when trying to assimilate remote sensing data into a well established model. Biases coming from the retrieval of the remotely sensed biophysical parameters might get introduced into the model. Due to the fact that only the initial conditions were fitted to the remote sensing data the reinitialization strategy is not as sensitive to parameter biases as, e.g., when forcing or updating model variables by the use of remote sensing data. In the presented approach it was also tried to account for the biases from dry matter and LAI estimation by weighting the parameters in respect to the achieved coefficients of determination from the remote sensing retrieval models.

5.8 Regional crop growth modelling

On the basis of mapped winter wheat fields in 2006, the soil map, slope, aspect and height information derived from the digital elevation model and the weather data in a first step regional yield estimation was done using the *automatic yield* approach. Overall, for the 435 mapped winter wheat fields, there was an average calculated final yield of 5655kg/ha. Using the multispectral data, without information about the ripening stage, the mean overall yield of all 435 mapped fields sums up to 7694kg/ha. In order to validate the findings, results were compared with the official statistics. The statistics of DESTATIS [1] list a total wheat yield for whole North Rhine Westphalia of 7840kg/ha. Compared to this the *automatic yield* approach underestimated the average final yield by 72%, which is a rather high bias. For the second used regional approach the bias sums up to only 1.86% underestimation. However, it is problematic to compare the modelling results with these statistics, due to the fact that the DESTATIS results cover a larger region. The final yield for the region of Meckenheim probably also lies above the mean for the whole North Rhine Westphalia, due to very good soils.
The described assimilation method, thus appears to be a promising technique to apply to crop growth models at regional scale, when no accurate in-situ information is available to run the model. The main critical aspect related to the presented approach is that the employed crop growth model needs to be calibrated, the genetic coefficients need to be adjusted, for the present wheat variety. The selected assimilation method provides an adjustment of management decisions, but not of the calibration properties (genetic coefficients). Concluding one can point out, that the illustrated approach can only be applied to agricultural areas relatively homogeneous in terms of crop variety, which is the case for the research region. Another critical aspect is, that the model also needs daily weather information for the whole region, which can become problematic when only one weather station is available for a large region.

5.9 Comparison to direct multi-sensoral yield modelling

Satellite remote sensing data have a long tradition in the field of crop yield modelling. The assumption hereby is, that the gained remote sensing information is strongly related to various crop parameters, e.g., biomass or LAI. These crop parameters are themselves, depending on the phenological stage, significantly related to final yield.
The combined yield retrieval can be formulated as:

$$Yield = A \times (NDVI - B) \times (\sigma^0 - C) \tag{5.1}$$

with Yield as the actual corn grain [kg/ha], NDVI as the used vegetation index, σ^0 as the backscatter intensity and the coefficients A, B and C (Heinzel, 2006; Liu et al., 2006). The calibration and validation of the model was done using the ground truth data from 2006 and a multispectral and ERS-2 image, acquired during stem elongation. Using this combination gave the best results, however the sampling amount was not that high, causing statistical limitations. The fitted coefficients for this modelling approach were: A=82397, B=-0.118 and C=-0.063

[1] http://www.destatis.de

with a R^2=0.98. When applying this model to the whole region, including the four monitored winter wheat fields, there was an underestimation of 3.03% for wheat field 1, and an underestimation for wheat field two by 11.64%. The other two fields were overestimated by 2.51% and 10.84%. When caluclating the total yield of these fields the trends of over- and underestimation cancel out, so that the overall yield estimation bias is only 0.02%. Concerning the total yield modelled for the whole region of Meckenheim it amounts to 7612kg/ha. Compared to the statistics of DESTATIS statistic (final average yield: 7840kg/ha), the modelled result for the region of Meckenheim features a bias of 2.91%. However, again the same regards as for the previously mentioned approach have to be made. It is problemtic to compare both, due to the fact that the DESTATIS results cover a larger region and probalby the final yield for the region of Meckenheim lies above the mean for the whole North Rhine Westphalia.
Compared to the findings of Liu et al. (2006) the result is reasonable. They achieved a $R = 0.83$ for final yield modelling using a Landsat TM and an ASAR (HH) scene during heading stage.

In Fig. 5.4 the final yield differences [kg/ha] between the just shown direct yield estimation method and the approach using the CERES model with multispectral data assimilation (without information about ripening stage). The majority of the fields feature a difference of +/- 1000kg/ha in final yield between both modelling approaches. Larger differences between both approaches can also be observed (red and darker green fields). When comparing these differences with the used soil map, it becomes apparent that especially on Stagnic soils the bias between both approaches was rather high. This is due to the fact, that the CERES-Wheat modelling actually takes into account different soils and the direct method does not.

When comparing the bias of both methods with the DESTATIS results, the CERES-Wheat final yield estimation slightly outperforms the direct method by being 1.05% preciser. However, as already mentioned both methods underestimated the final yield in comparison to the DESTATIS results. Which is critical, because the actual final yield within the region of Meckenheim might even be higher, as noted by the DESTATIS statistic, due to very good soils.
Another aspect, which has to be considered is, that using the direct yield estimation method is not as complicated as the use of the CERES-Wheat model, by still providing reasonable results. The method has the disadvantage of not being based on physical, physiological and biochemical laws. This can become problematic, when trying to transfer the findings in time or space. As it is also not based on other ground truth data like weather or management information, the approach has problems in situations of sudden weather anomalies occurring after the remote sensing data acquisition.

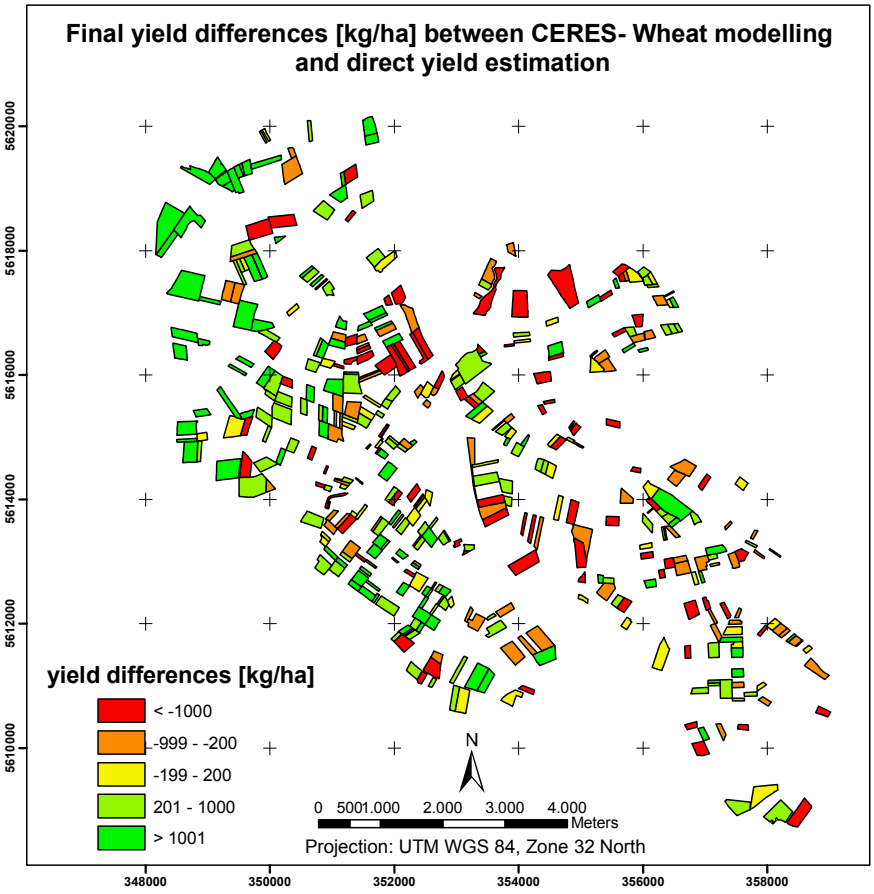

Figure 5.4: Final yield differences between modelling using the CERES-Wheat model and direct final yield estimation [kg/ha].

5.10 Conclusion of the crop growth modelling

Within this last application section the following question and assumptions were assessed:

Is it possible to integrate important crop variables gained from multi-source data into a dynamic crop growth model in order to increase the final yield estimation accuracy?

Challenges hereby were:

- to assimilate biophysical parameter maps retrieved from different remote sensing data into a crop growth model in order to improve the accuracy of wheat yield predictions
- to analyse the effect of assimilating separately or synergistically SAR and multispectral time-series information
- to analyse the effect of acquisition time and frequency of the assimilated multispectral data
- to assess whether time gaps in the multispectral data series can be filled by SAR informations
- to transfer the findings of the point-based crop growth model for regional final yield prediction
- to analyse the performance of the suggested approach in comparison to a simple direct yield estimation method

The previous analysis has shown that overall the integration of additional information, as remotely sensed biophysical parameters, improved the final yield estimation accuracy of the CERES-Wheat model. When using an automatic setting the bias in final yield estimation was around 6.6%, this improved significantly for all assimilated time-series. The best results for this first analysis were achieved by integrating the whole available multi-sensoral time-series (SAR and multispectral information), ending up with a bias of only 0.56% in final yield for the four intensively monitored winter wheat fields.
It was further assessed that the acquisition time and amount of multispectral information also has an impact on the final yield modelling. Especially, during rather early phenological stages (stem elongation) the multispectral information was very important. However, in the case of missing multispectral information during this stage additional SAR information were not able to fill the multispectral time gap. As also for missing multispectral information during flowering the modelling bias even increased when also regarding the SAR information. Concerning the multispectral information during the ripening stage, it actually had a negative influence on the results, no matter if the SAR information was also additionally considered.
The transfer of the actually point-based model to regional modelling worked satisfactory, whereby a really statistically based assessment was not possible. But it appears that, with the used the calibrated genetic coefficients and the remotely sensed information, regional modelling performed sound. Also in comparison to an automatic parameter setting of the CERES-Wheat model, or a direct multi-sensoral regional yield estimation approach, the suggested approach outperformed the others.

6 GENERAL DISCUSSION AND CONCLUSIONS

Population growth, increasing living standards, globalization, and global climate change increase the pressure on the limited natural resources. Natural resources have to be carefully managed in order to ensure their sustainable use. Food security and safety, environmental friendly production technologies, sustained use of renewable and other resources are important issues. Remote sensing technologies can help monitoring the landscape, assessing changes as well as predicting final yield. Especially, the large spatial coverage and multi-sensoral data are of advantage. Crop growth modelling is also an important technology in this context. The models can help to understand the complex relationships and interactions between the different systems, e.g., plants, water, atmosphere. But they can also be used for final yield prediction, or as decision support systems, assessing, e.g., the impact of climate change, or herbicide, fungicide and pesticide applications, also within the context of precision farming and the minimisation of the application amount. Combining crop growth modelling and remote sensing technologies enables to take advantage of each and allows to apply the actually point-based crop growth models into a more spatial context.
In the presented study the scope and constraints of biophysical parameter derivation from multi-sensoral remote sensing data and their integration into a crop growth model for regional yield assessment were investigated.

6.1 Gained insights

Biophysical parameter estimation from remotely sensed data is a challenging and important topic, especially for regional applications. The biophysical parameters assessed were fresh and dry biomass, plant water content and leaf area index of winter wheat, as well as the important phenological stage heading/flowering. Crop growth models can be successfully applied if their limitations are taken into account and if the huge amount of required input data are provided. The combination of both leads to better yield predictions, and when used for decision support to improved management practices and reduced environmental impact. Whereby, still random or unexpected events, e.g., pest or disease infections or extreme weather conditions or flooding, may cause the model to fail or simulation results to differ from reality, if remotely sensed information, covering that time period, are not available. Here lies the advantage of using multi-sensoral data and especially the weather-independent SAR data.
In chapter 1, the thesis started with the formulation of a number of research objectives and hypotheses, which were afterwards examined and analysed. The first general objective was to assess the possibilities of deriving biophysical parameters from different remote sensing data, with an appropriate accuracy and statistically validated. Therefore, either single source remote

sensing data or multi-sensoral data were analysed. One important feature was to use operational models, which were easy to invert enabling regional applications. The used models should be in agreement with theoretical findings, but not as complex as deterministic models. The other general objective was the integration of the derived crop variables into a dynamic crop growth model in order to increase the final yield simulation accuracy and enable regional applications.
In chapter 2 the complex theoretical background was introduced. The first concerns were considerations about multispectral remote sensing data and vegetation monitoring, multi-temporal vegetation signatures and intercalibration issues between vegetation indices derived from different multispectral sensors. In the following theoretical subsection, issues dealing with SAR system features and vegetation target features, as well as their multi-temporal changes were introduced. After this more general introduction of the remote sensing methods, different retrieval options of vegetation features from these data were introduced. Important issues were the regional and operational use of the suggested methods and the issue of multi-sensoral derivation methods, using the advantages of each remote sensing system. Within the last theoretical section general concerns of the phenological development from the assessed crop *winter wheat* as well as its management, crop growth modelling and final yield estimation, and finally the possibilities of assimilating remotely sensed information into these models, were introduced.
After presenting the complex theoretical background, the used data were introduced in chapter 3. In the first section the complex ground truth campaigns were summarized. These campaigns were accomplished in weekly intervals during the vegetation periods 2005 and 2006. Important issues were the data sampling set up and the different measurement methods of the monitored variables, as well as their statistical analysis. The following section in chapter 3 dealt with the various multispectral data and the used SAR data. The important issues were their multi-temporal coverage and the different necessary preprocessing steps. An accurate preprocessing is especially important when working with multi-temporal and multi-sensoral data, due to comparability issues. Concerning the multispectral data one feature was the accurate atmospherical correction, but also an accurate geometrical correction of all used data.
In chapter 4 all the previously described issues and data were considered for the actual applications, analysis and assessments of the first work hypotheses.
The biophysical estimation example (section 4.1) focussed on the extraction of the desired biophysical parameters from multispectral remote sensing data. Different empirical modelling methods using different vegetation indices were assessed. It became evident, that when working with different multispectral sensors an intercalibration of the different calculated vegetation indices is an necessary additional processing step, due to differences in the respective relative response functions. For the intercalibration process an empirically based method was proposed, developed and validated on the basis of simulated multispectral remote sensing sensors. After intercalibration the vegetation index differences, introduced by different relative response functions, decreased significantly. Differences introduced by other sensor characteristics, e.g., different spatial resolution or acquisition geometry aspects, were still present in the data. Overall, even after intercalibration, the proposed empirical models did not derive the wbiophysical parameters with the desired accuracy. In a next step the retrieval performance of the semi-empirical CLAIR model, actually designed for LAI retrieval, was assessed. Findings indicated that the model can also be used for the extraction of other biophysical parameters and outperforms the other previously evaluated empirical models. It generally was necessary to intercalibrate the used vegetation index WDVI to one reference sensor and to divide the assessed time-series into before and after heading/flowering. Findings then also enabled regional parameter estimation.

Further analyses (section 4.2) addressed the parameter retrieval from the SAR time-series. One important aspect were the performances of the different used polarizations (VV, HH, HV, HH/HV ratio). In a first step modelling performances were assessed by using simple empirical regressions. Due to the complexity of the target/sensor interactions it was found that multiple regression, taking into account the local incidence angle, roughness features (vegetation and soil) and precipitation sums, as an indicator of the soil moisture, outperformed the simple regressions. Because of regional parameter retrieval and the problem of regionally assessing roughness features, these variables were later left out. Even the reduced multiple regression exceeded the simple regressions. The performance of the semi-empirical Water Cloud Model for biophysical parameter extraction was also assessed, once with the additional parameter vegetation roughness, and once without. Generally, the reduced multiple regressions outperformed both Water Cloud Model types in the retrieval performance. Concerning the potential of the different polarizations and incidence angles for retrieval purposes the findings were a bit contradictory. Reasonable results for the VV polarization could only be achieved for the retrieval of fresh biomass and the plant water content. For the ASAR sensor modelling fresh biomass and LAI using the HV polarization or the dry biomass using the ratio was appropriate. No valid results could be achieved for the ASAR modelling and the plant water content. It was not possible to make an overall assumption about the most appropriate polarization to use. Concerning the incidence angle the first simple analysis showed that higher incidence angles like the ASAR swath 6 outranges lower incidence angles, due to a smaller soil signal amount in the backscattered signal. The heading/flowering date extraction on the basis of VV time-series analysis worked sound for all assessed fields and years. All findings could be applied for regional parameter estimation. In comparison to the multispectral retrieval performances SAR findings feature a higher bias in the parameter estimation as modelling with the use of the CLAIR model. In section 4.3. multi-sensoral biophysical parameter estimation methods were analysed. These methods take advantage of the fact that multispectral and SAR systems interact in different ways with the monitored targets. In a first step combined modelling was realized for all polarizations using a simple linear approach for the combination of the NDVI and the backscatter. Overall the findings were not reasonable for the assessed parameters and polarizations. Only for the NDVI and ASAR ratio the combined fresh biomass retrieval method slightly outperformed the single sensor modelling. In the next steps more complex combination methods, also using different weighting factors for the respective parameters and data types were applied. The most appropriate combined method was the CLAIR model combined with multiple regression for the ERS-2 data. In comparison to the single sensor modelling approaches, there was a clear improvement of the results, especially during late phenological stages. Only for the parameter fresh biomass there was an obvious trend of parameter underestimation.
The remotely sensed information was then finally assimilated into the CERES-Wheat crop growth model (chapter 5). Therefore first of all a sensitivity analysis of the CERES-Wheat input parameters was necessary, in order to assess, whether there are input conditions, strongly affecting the crop growth. Seeding and fertilization date appeared to have the strongest impact on the final yield and crop growth (section 5.3). These two parameters were then re-initialized by the use of the different remotely sensed variables. When assimilating the different time-series (first approach) into the model, it became obvious that the best performance could be achieved when the full remote sensing data time-series, containing SAR and multispectral data was used (section 5.4). Concerning the performance of the multispectral data assimilation, the performance actually increased when leaving out the LAI and dry matter maps from late phenological stages (ripening). The achieved final yield estimation quality was then as sound as using the full multi-sensoral data series (section 5.5).

When trying to fill the gaps in the multispectral time-series by using SAR information, the precision for the first two approaches (no multispectral information from stem elongation or flowering) even decreased the final yield estimation accuracy. When leaving out multispectral information during ripening the result was comparable to the achievement without considering SAR information. The SAR information was not able to fill multispectral time gaps during important phenological stages (section 5.6). Overall, modelling by the use of remotely sensed data assimilation outperformed the CERES-Wheat modelling with automatic input condition setting. For regional application the suggested method outperformed the automatic alternative and another evaluated method of direct yield estimation using one ERS-2 acquisition and the NDVI (section 5.9).

6.2 Validation of the hypotheses

In chapter 1 a number of objectives and questions to be examined were formulated. In the foregoing section the different application examples were briefly summarized. Now the hypotheses will either be accepted or rejected, with a short reasoning. The short forms and numbers in the brackets, introduced below, are used in the following table (Tab. 6.1), showing the validation of the different hypotheses in three different categories.
The examined hypotheses were:

Is it possible to derive the values of important crop variables from various remote sensing data (1 parameter derivation)?

The answers to the following challenges are finally given in Tab. 6.1:

- to intercalibrate vegetation indices of different multispectral sensors in order to correct the biases, introduced by differences in the relative spectral response functions (1.1 intercalibration)
- to evaluate the performance of the CLAIR model for the retrieval of different biophysical parameters (1.2 CLAIR model)
- to analyse the potential of different polarizations and incidence angles for the retrieval of different biophysical parameters (1.3 SAR system aspects)
- to analyse the performance of the Water Cloud model and multiple regressions using different SAR data and the additional the variable *vegetation roughness* (1.4 vegetation roughness)
- to extract important phenological events such as heading/flowering from remote sensing information for the adjustment of the retrieval models and the simulation model (1.5 heading/flowering)

- to analyse the potential of combining multispectral and SAR data for the derivation of biophysical parameters (1.6 combined modelling)

- to invert the used models for regional biophysical parameter extraction (1.7 regional application)

These derived biophysical parameters were then assimilated into a crop growth model for final yield estimation, whereby the answers to the following questions are also given in Tab. 6.1:

Is it possible to integrate important crop variables gained from multi-source data into a dynamic crop growth model in order to increase the final yield estimation accuracy (2 final yield estimation)?

Challenges hereby were:

- to assimilate biophysical parameter maps retrieved from different remote sensing data into a crop growth model in order to improve the accuracy of wheat yield predictions (2.1 assimilation)

- to analyse the effect of assimilating separately or synergistically SAR and multispectral time-series information (2.2 data sources)

- to analyse the effect of acquisition time and frequency of the assimilated multispectral data (2.3 time and frequency)

- to assess whether time gaps in the multispectral data series can be filled by SAR informations (2.4 SAR filling)

- to transfer the findings of the point-based crop growth model for regional final yield prediction (2.5 point-based to regional)

- to analyse the performance of the suggested approach in comparison to a simple direct yield estimation method (2.6 comparison)

Overall, both main questions were satisfactorily fulfilled and most sub-hypotheses can be accepted. It was possible to derive statistically valid the biophysical parameters from the remote sensing data (question 1).
Hypothesis 1.1 (intercalibration) can also be accepted as the suggested intercalibration method performed well for the evaluated vegetation indices (NDVI, WDVI) and reduced the sensor differences significantly. This also had a positive impact on the biophysical parameter estimation accuracy.
Hypothesis 1.2 (the biophysical parameter modelling by the use of the CLAIR model) can only be partly accepted, due to the fact that modelling the dry matter biomass during late phenological stages featured a high trend of parameter underestimation and only a R^2 of 0.56. For all other assessed parameters and times the model performed valid, hence the CLAIR model, actually designed for LAI retrieval, can also be used for the derivation of the other analysed parameters.

Table 6.1: Hypotheses acceptance or rejection (short forms are explained in the text above).

hypotheses	accepted	half accepted	rejected
1 parameter derivation		X	
1.1 intercalibration	X		
1.2 CLAIR model		X	
1.3 SAR system aspects		X	
1.4 vegetation roughness	X		
1.5 heading/flowering	X		
1.6 combined modelling		X	
1.7 regional application	X		
2 final yield estimation	X		
2.1 assimilation	X		
2.2 data sources	X		
2.3 time and frequency	X		
2.4 SAR filling			X
2.5 point-based to regional	X		
2.6 comparison	X		

(accepted = hypotheses fully valid; half accepted= there were some problems; rejected= hypothesis has to be rejected.)

Hypothesis 1.3 the analysis of the potential of different SAR incidence angles and polarizations can also only be partly accepted. First analysis showed that higher incidence angles (ASAR swath 6) feature a high potential in biophysical parameter estimation, however due to the small amount of ASAR data with different incidence angles, the hypothesis could not be validated statistically. Concerning the different evaluated polarizations (VV, HH, HV, HH/HV (ratio)) no clear conclusion can be drawn; there was no polarization outperforming all others for all evaluated parameters. For the parameters fresh biomass and LAI the HV polarization outperformed the others, for the plant water content best retrieval results were achieved using the VV polarization and for the dry biomass the ratio (HH/HV) outperformed the others.
The 1.4 hypothesis concerned the integration of the parameter vegetation roughness for the retrieval using SAR data. For all tested retrieval algorithms the additional use of this variable improved the modelling performances significantly, especially for the parameter fresh biomass.
Heading/flowering date extraction (hypothesis 1.5) was fully achieved by the use of the ERS-2 (VV) time-series. The findings could also be used for regional heading/flowering date estimation.
The multi-sensoral parameter estimation (hypothesis 1.6 combined modelling) was successful for most of the tested times and parameters, outperforming the previous single sensor estimations. However, the findings for the parameter fresh biomass, especially during early phenological stages were slightly poorer than the performance just using the CLAIR model (multispectral data).
The last hypothesis concerning the biophysical parameter estimation (1.7 regional application) can be accepted. The chosen models only needed input parameters, which could be easily gained for large areas and the model inversion also featured no problems. Only the roughness parameters (soil and vegetation) could not be measured for larger regions, hence the modelling approaches using these variables could not be used for regional modelling.

The second main question (2 final yield estimation) dealing with the improvement of final yield estimation by assimilating remotely sensed information into the CERES-Wheat model was also successfully answered. For all assimilated data series there was a clear improvement in the final yield estimation in comparison to automatically input condition setting of the CERES-Wheat model or direct final yield estimation. Hence, the hypothesis 2.1 (assimilation) can also be accepted.
Concerning the different data sources (hypothesis 2.2), actually the multispectral data series and the dense data series using all remotely sensed information in a combined data series (SAR and multispectral) outperformed the other options evaluated. The combined data series slightly outperformed the full multispectral series.
The analysis of different time gaps in the multispectral data series (2.3 time and frequency) was also analysed successfully. Time gaps during early phenological stages (stem elongation) increased the bias in the final yield estimation. Gaps after this phenological stage, actually improved the modelling performance, hence a too dense data series can also have a negative effect on the final yield estimation, probably due to the biases in the remotely sensed information.
The hypothesis 2.4 (SAR filling) dealt with the question, if data gaps in the multispectral data series can be filled by using information from SAR data. Within this study the hypothesis actually has to be rejected. The modelling performance when filling the time gaps was either the same or even featured a decreased performance, as in the cause of no available multispectral information during stem elongation or flowering. One can actually state, that the SAR information is not appropriate for filling missing time gaps in the multispectral time-series. It more appears, that the modelling performance even decreased in regard to only using a multispectral data series, containing time gaps of distinctive phenological periods.
Using the remotely sensed information (2.5 point-based to regional) for regional crop growth with the CERES-Wheat model worked sound and delivered reasonable results. Problematic is that, the results could not be evaluated finally, due to the fact that actual final yield data were not available for each field within the analysed region. The used DESTATIS data for the comparison might not be that appropriate to use. It covers a larger region than the assessed region of Meckenheim.
Compared with other methods (2.6 comparison) the suggested method using the CERES-Wheat model with assimilated multispectral data (no information about ripening) outperformed the option of automatic input parameter setting of the CERES-Wheat model and another tested method of direct yield estimation using the vegetation index NDVI and one SAR acquisition.

6.3 OUTLOOK – future research and challenges

Finally, an outlook will be given highlighting the possibilities of future research and challenges.

New satellite sensors and application fields
Future satellite missions and just recently started satellites, e.g., TerraSAR-X, bring new challenges and options concerning all remote sensing fields. The new SAR satellites possess new features. TerraSAR-X is a X-band sensor with a higher spatial resolution and the ability of acquiring in a full polarimetric mode. The Japanese ALOS satellite is a L-band system, also with the ability of full polarimetric data acquisition. First acquired data of this satellite have already shown the high potential. Additionally, there are also future missions as Radarsat-2 or the Sentinel mission, both C-band systems.
These new systems have a high potential of synergetic use, e.g., combining the different frequencies. Also the higher spatial resolution and the ability of some sensors providing data in full polarimetric mode, feature new possibilities.
New multispectral or hyperspectral satellites (RapidEye or EnMap) with high spatial resolution and high repetition rate (RapidEye), increase the possibilities of monitoring high dynamic processes as plant growth.
Overall, each system features a new potential itself, but also the combined use of the different sensor systems will be an important research topic in the future.

Biophysical parameter retrieval
Within the discussions during the BioGeoSAR 2007 in Italy, it became clear that there is a big need for new developments concerning the modelling of biophysical parameters from crops. The last achievements, models and detailed analysis are rather old from Attema (Attema & Ulaby, 1978) or Ulaby (Ulaby et al., 1986). New approaches have to stay rather simple, making it possible to invert the respective models for regional estimations and operational use. However, they should still be in agreement with physical based models. A challenge in this context is to take into account roughness and soil moisture features, maybe by using generalized assumptions or by coupling with other models. Models also have to take into account new SAR satellite missions, with other frequencies, e.g., L-band and or X-band, and the possibility of full polarimetric data acquisition.
Due to new multispectral and hyperspectral satellites, there will also be a need for the adjustment of the semi-empirical models to these data. The issue of intercalibration between different satellites will be a topic in the future, due to the huge variety of different multispectral/hyperspectral satellites, each featuring different relative response functions.
For all new sensor systems the topic of different spatial resolution and higher repetition rate will bring new challenges and opportunities for biophysical parameter retrieval.
Considering the combined use of different satellite systems, developing new methods and models also for other crop characteristics, e.g., plant stress, different phenological stages, will be a challenging topic in the future. The presented thesis can be seen as a first basis for synergetic biophysical parameter retrieval. Other fusion techniques might also have to be considered, like the decision based fusion. Fusing maybe the individual mono-sensorally retrieved biophysical parameters in a later processing step, by using different weighting factors.

Crop growth modelling
Crop growth modelling has experienced a great increase within the last years, due to new technologies and a better understanding of the underlying processes and interactions. A great challenge still is to apply the point-based models to a spatial scale by running them on a grid of pixels or for a set of polygons, resulting from image analysis and classification procedures. This spatial analysis would bring the possibility of analysing specific spatial patterns of larger regions or in the context of precision farming. Also within the context of decision support systems or the impact of climate change on crops, there is still scientific research demand.

Data assimilation
It may be attractive to derive additional crop state variables from remote sensing data and integrate those in crop growth simulation models, especially if the observation frequency is high. Regarding this, the topic of different assimilation methods should also be taken into account, as well as the use of filters like the Kalman filter and the handling of biases in the remotely sensed parameters. This is especially important, if multi-source data are used containing different error sources. As well as when using e.g., classification algorithms for the derivation of land cover maps, which are integrated into the models for regional modelling purposes.

The presented thesis can be seen as a first basis, highlighting the possibilities of new satellite sensors. Especially, in regard to multi-sensoral analysis, fusing multispectral and SAR information, appears to be a promising technique for the future. It is expected that in the context of biophysical parameter retrieval the suggested approaches can lead to an improvement of the estimation accuracy. Exacter biophysical parameter estimation will then also have a positive impact on the crop growth modelling performances. As shown in the thesis, remote sensing information can help re-initializing the models and enable spatial modelling.

Bibliography

Albertz, J. (2001). Einführung in die Fernerkundung. Grundlagen der Interpretation von Luft- und Satellitenbildern. Wissenschaftliche Buchgesellschaft.

Amodeo, G., de Matthaeis, P., Ferrazzoli, P., Paloscia, S., Pampaloni, P., Schiavon, G., Sigismondi, S., & Solimini, D. (1996). Monitoring vegetation features with multi-temporal sar data. In Proc. Int. Geosci. Remote Sensing Symp. (IGARSS '96), Lincoln, Nebraska.

Arnold, A. (Ed.). (1997). Allgemeine Agrargeographie. Klett-Perthes.

Attema, E. & Ulaby, F. (1978). Vegetation modelled as a water cloud. Radio Science, 13(2), 357–364.

Atzberger, C. (2003). Möglichkeiten und grenzen der fernerkundlichen bestimmung biophysikalischer vegetationsparameter mittels physikalisch basierter reflexionsmodelle. PFG, 1, 51–61.

Bach, H. & Mauser, W. (2003). Methods and examples for remote sensing data assimilation in land surface process modeling. IEEE Trans. Geosci. Remote Sensing, 41(7), 1629–1637.

Bannayan, M., Crout, N., & Hoogenboom, G. (2003). Application of the ceres-wheat model for within-season prediction of winter wheat yield in the united kingdom. Agron. J., 95(1), 114–125.

Baret, F. & Guyot, G. (1991). Potentialy and limits of vegetation indices for lai and apar assessment. Remote Sens. Environ., 35(2-3), 161–173.

Bastiaanssen, W. (1998). Remote sensing in water resources management: The state of the art. Technical report, International Water Management Institute. Colombo. Sri Lanka.

Bauer, M., Daughtry, C., Biehl, L., Kanemasu, E., & Hall, F. (1986). Field spectroscopy of agricultural crops. IEEE Trans. Geosci. Remote Sensing, GE-24(1), 65–75.

Borgeaud, M., Noll, J., & Bellini, A. (1995). On the use of ers-1 multi-temporal sar data for agricultural applications. In International Geoscience and Remote Sensing Symposia, IGARSS.

Bouman, B. (1992). Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. Int. J. Remote Sensing, 13(14), 2565–2581.

Bouman, B. & van Kasteren, H. (1990). Ground-based x-band (3cmwave) radar backscattering of agricultural crops. ii: Wheat, barley and oats; the impact of canopy structure. Remote Sens. Environ., 34(2), 107–118.

Braga, R. (2000). Predicting the spatial patterns of grain yield under water limiting conditions. PhD thesis, Agricultural and Biological Engineering Department, University Florida, USA.

Brakke, T., Kanemasu, E., & Steiner, J. (1981). Microwave radar response to canopy moisture, leaf-area index and dry weight of wheat, corn, and sorghum. Remote Sens. Environ., 11, 207–220.

Bréda, N. (2003). Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. J. Exp. Bot., 54(392), 2403–2417.

Brisco, B. & Brown, R. (1998). Principles & Applications of IMAGING RADAR, Manual of Remote Sensing, Third Edition, Volume 2. John Wiley and Sons, Inc.

Cassman, K. (1999). Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. Proc. Natl Acad. Sci., 96, 5952–5959.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., Aubinet, M., & Buchmann, N. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. Nature, 437, 529–533.

Clevers, J. (1988). The derivation of a simplified reflectance model for the estimation of leaf area index. Remote Sens. Environ., 25(1), 53–69.

Clevers, J. (1989). The application of a weighted infrared-red vegetation index for estimating leaf area index by correcting for soil moisture. Remote Sens. Environ., 29(1), 25–37.

Clevers, J. & Van Leeuwen, H. (1996). Combined use of optical and microwave remote sensing data for crop growth monitoring. Remote Sensing Environ., 56(1), 42–51.

Clevers, J., Vonder, O.W. an Jongschaap, R., Desprats, J. F., King, C., Prevot, L., & Bruguier, N. (2002). Using spot data for calibrating a wheat growth model under mediterranean conditions. Agronomie, 22, 687–694.

Colombo, R., Bellingeri, D., Fasolini, D., & Marino, C. (2003). Retrieval of leaf area index in different vegetation types using high resolution satellite data. Remote Sens. Environ., 86(1), 120–131.

Dactu, M., Seidel, K., D´Elia, S., & Marchetti, P. (2002). Knowledge-driven information mining in remote-sensing image archieves. ESA bulletin, 3, 26–33.

Delecolle, R., Maas, S., Guerif, M., & Baret, F. (1992). Remote sensing and crop production models: present trends. ISPRS J. Photogramm. Remote Sens., 47(2-3), 145–161.

Dente, L., Rinaldi, M., Mattia, F., & Satalino, G. (2004). On the assimilation of c-band radar data into ceres-wheat model. In International Geoscience and Remote Sensing Symposia, IGARSS.

Dente, L., Satalino, G., Mattia, F., & Rinaldi, M. (2007). Assimilation of leaf area index derived from asar and meris data into ceres-wheat model to map wheat yield. Remote Sens. Environ. in press., doi:10.1016/j.rse.2007.05.023, 1–13.

Dobson, M., Ulaby, F., & Pierce, L. (1995). Land-cover classification and estimation of terrain attributes using synthetic aperture radar. Remote Sens. Environ., 51(1), 199–214.

Doraiswamy, P., Moulin, S., Cook, P., & Stern, A. (2003). Crop yield assessment from remote sensing. Photogrammetric Engineering & Remote Sensing, 69(6), 665–674.

Dorigo, W., Zurita-Milla, R., de Wit, A., Brazile, J., Singh, R., & Schaepman, M. (2007). A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. Int. J. Appl. Earth Observ. Geoinform., 9(2), 165–193.

Dworschak, M., Schrey, H., & Schulte-Kellinghaus, S. (2001). Allgemeine informationen zur bodenkarte 1 : 50 000. - 55 s. Technical report, Arbeitsgruppe BK50, Krefeld (Geol. Dienst Nordrh.-Westf.).

Fang, W., Chen, J., & Shi, P. (2005). Variability of the phenological stages of winter wheat in the north china plain with noaa/avhrr ndvi data (1982-2000). In International Geoscience and Remote Sensing Symposia, IGARSS.

Ferrazzoli, P., Wigneron, J., Guerriero, L., & Chanzy, A. (2000). Multifrequency emission of wheat: Modeling and applications. IEEE Trans. Geosci. Remote Sensing, 38(6), 2598–2607.

Foley, J., DeFries, R., Asner, G., Barford, C., Bonan, G., Carpenter, S., & Chapin, F. (2005). Global consequences of land use. Science, 309, 570–573.

Fontana, D. C., Potgieter, A. B., & Apan, A. (2005). Assessing the relationship between shire winter crop yield and multi-temporal modis ndvi and evi images. In SSC2005 Spatial Sciences Institute Biennial Conference.

Franke, J., Heinzel, V., & Menz, G. (2006). Assessment of ndvi-differences caused by sensor-specific relative spectral response functions. In International Geoscience and Remote Sensing Symposia, IGARSS.

Gao, B. (2000). A practical method for simulating avhrr-consistent ndvi data series using narrow modis channels in the 500-1000nm sprectral range. IEEE Trans. Geosci. Remote Sensing, 38(4), 1969–1975.

Geisler, G. (1970). Pflanzenbau in Stichworten. I. Die Kulturpflanzen. Ferdinand Hirt.

Gittelson, A. & Kaufman, Y. (1998). Modis ndvi optimization to fit the avhrr data series-spectral considerations. Remote Sens. Environ., 66(3), 343–350.

Graham, A. & Harris, R. (2002). Estimating crop and waveband specific water cloud model parameters using a theoretical backscattering model. Int. J. Remote Sensing, 23(23), 5129–5133.

Guissard, V., Lucau-Danila, C., & Defourny, P. (2005). Crop specific lai retrieval using optical and radar satellite data for regional crop growth monitoring and modelling. In SPIE, Society of Photo-Optical Instrumentation Engineers, Brügge, Belgium.

Hamacher, M. (2000). Zustandserfassung von Wintergetreide und Zuckerrüben während des Wachstums mit Hilfe des C-Band Radars der ERS-1 und ERS-2 Satelliten. PhD thesis, University Bonn, Germany.

Heinzel, V. (2006). Synergetic use of optical and ers-2 data for crop yield retrieval. In 2nd Workshop of the EARSeL Special Interest Group on Land Use and Land Cover.

Heinzel, V., Franke, J., & Menz, G. (2006). Assessment of cross-sensor ndvi-variations caused by spectral band characteristics. In SPIE.

Heinzel, V., Franke, J., & Menz, G. (2007). Assessment of differences in multi-sensoral remote sensing imageries caused by discrepancies in the relative spectral response functions. In Internationales Wissenschaftliches Kolloquium, IWK.

Heinzel, V., Waske, B., Braun, M., & Menz, G. (2007). Remote sensing data assimilation for regional crop growth modelling. In International Geoscience and Remote Sensing Symposia, IGARSS.

Henderson, F. & Lewis, A. (1998). Principles and applications of imaging radar (Third Edition). Manual of Remote Sensing vol. 2 (1998). American Society of Photogrammetric Engineering and Remote Sensing.

HGCA (1998). The wheat growth guide. Technical report, Home-Grown Cereals AUthority. Caledonia House. London.

HGCA (2000). Optimum winter wheat plant population. Technical report, Home-Grown Cereals AUthority. Caledonia House. London.

HGCA (2002). Precision farming of cereals, practical guidelines and crop nutrition. Technical report, Home-Grown Cereals AUthority. Caledonia House. London.

HGCA (2004). Assessment of sensor-based technologies for monitoring crop growth and development in cereals. Technical report, Home-Grown Cereals AUthority. Caledonia House. London.

Hoyt, P. & Bradfield, R. (1962). Effect of varying leaf area by partial of the meter by partial defoliation and plant density on dry matter production in corn. Agron. J., 54, 523–525.

Huete, A. (1988). A soil-adjusted vegetation index (savi). Remote Sens. Environ., 25(3), 295–309.

Hunt, L., Pararajasingham, S., Jones, J., Hoogenboom, G., Imamura, D., & Ogoshi, R. (1993). Gencalc - software to facilitate the use of cropmodels for analysing field experiments. Agron. J., 85, 1090–1094.

Jackson, R. (1983). Spectral indices in n-space. Remote Sens. Environ., 13(5), 409–421.

Jacquemoud, S. & Baret, F. (1990). Prospect: a model of leaf optical properties spectra. Remote Sens. Environ., 34(2), 75–91.

Jensen, J. (2000). Remote Sensing of the Environment: An Earth Resource Perspective. Prentice Hall: Saddle River, N.J.

Jensen, J. (Ed.). (2007). Remote Sensing of the Environment: An Earth Resource. Prentice Hall.

Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., & Ritchie, J. (2003). The dssat cropping system model. Europ. J. Agronomy, 18(3-4), 235–265.

Jones, J., Keating, B., & Porter, C. (2001). Approaches to modular model development. Agricultural Systems, 70(2-3), 421–443.

Jongschaap, R. (2006). Integrating crop growth simulation and remote sensing to improve resource use efficiency in farming systems. PhD thesis, University Wageningen. Neatherland.

Jongschaap, R. & Schouten, L. (2005). Predicting wheat production at regional scale by integration of remote sensing data with a simulation model. Agron. Sustain. Dev., 25, 481–489.

Kerekes, J. P. & Landgrebe, D. A. (1989). Simulation of optical remote sensing systems. IEEE Trans. Geosci. Remote Sensing, 27(6), 762–771.

Kühbauch, W. & Hawlitschka, S. (2003). Remote sensing - a future technology in precision farming. In Proceedings of the Workshop on POLinSAR - Applications of SAR Polarimetry and Polarimetric Interferometry (ESA SP-529).

Lancashire, P., Bleiholder, H., Langeluddecke, P., Stauss, R., van den Boom, T., Weber, E., & Witzen-Berger, A. (1991). An uniform decimal code for growth stages of crops and weeds. Ann. appl. Biol., 119, 516–601.

Laur, H. (1998). Calibration of the ers sar pri data product. earthnet online.

Lehmann, M.; Block, T. & Pätzold, S. (1996). Auswirkungen von Pflanzenschutzmitteln im Obstbau auf die Ertragsleitsung und Qualität von Ernteprodukten sowie auf die Nützlingspopulationen und Eigentschaften der Böden. Institut für Obstbau und Gemüsen; Universität Bonn.

Leser, H. (1998). Wörterbuch Allgemeine Geographie. Diercke.

Lillesand, T. & Kiefer, R. (2000). Remote Sensing and Image Interpretation. John Wiley & Sons, Inc.

Liu, L., Wang, J., Bao, Y., Huang, W., Ma, Z., & Zhao, C. (2006). Predicting winter wheat condition, grain yield and protein content using multi-temporal envisat-asar and landsat tm satellite images. Int. J. Remote Sensing, 27(4), 737–753.

Lopes, A., Touzi, R., & Nezry, E. (1990). Adaptive speckle filters and scene heterogeneity. IEEE Trans. Geosci. Remote Sensing, 28(6), 992–1000.

Maas, S. (1988a). Use of remotely-sensed information in agricultural crop growth models. Ecological Modelling, 41(3-4), 247–268.

Maas, S. (1988b). Using satellite data to improve model estimates of crop yield. Agron. J., 80, 655–662.

Mather, P. (1999). Computer Processing of Remotely-Sensed Images - An Introduction. Wiley.

Mattia, F., Dente, L., G., S., & Le Toan, T. (2004). Sensitivity of asar ap data to wheat crop parameters. In Proc. of the 2004 Envisat & ERS Symposium, Salzburg, Austria.

Mattia, F., Le Toan, T., Picard, G., Posa, F., DÀlessio, A., Notarnicola, C. amd Gatti, A., Rinaldi, M., Satalino, G., & Pasquariello, G. (2003). Multitemporal c-band radar measurements on wheat fields. IEEE Trans. Geosci. Remote Sensing, 41(7), 1551–1558.

Mavromatis, T., Boote, K., Jones, J., Wilkenson, G., & Hoogenboom, G. (2002). Repeatability of model genetic coefficients derived from soybean performance trials cross different states. Crop Sci., 42(1), 76–89.

McConnell, P. (1995). The agricultural notebook. Blackwell Science, London.

Miura, T. amd Huete, A. & Yoshioka, H. (2006). An empirical investigation of cross-sensor relationships of ndvi and red/near-infrared reflectance using eo-1 hyperion data. Remote Sens. Environ., 100(2), 223–236.

Monteith, J. (1977). Climate and the efficiency of crop production in britain. Phil. Trans. R. Soc. Lond. B, 281(980), 277–294.

Moore, R. (1970). Ground echo. Radar Handbook. McGraw-Hill.

Moulin, S., Bondeau, A., & Delecolle, R. (1998). Combining agricultural crop models and satellite observations: from field to regional scales. Int. J. Remote Sensing, 19(6), 1021–1036.

Norman, J. & Campbell, G. (1989). Canopy structure, Plant Physiological Ecology: Field methods and instrumentation. Chapman & Hall, London and New York.

Oliver, C. & Quegan, S. (2004). Understanding Synthetic Aperture Radar Images. Scitech. Publishing. Inc.

Otter-Nacke, S., Godwin, D., & Ritchie, J. (1986). Testing and validating the ceres-wheat model in diverse environments. Technical report, Publ. YM-15-00407 JSC 20244. LBJ Space Cent. Houston, TX: Earth Res. Appl. Div.

Patel, H. & Shekh, A. (2005). Sensitivity analysis of ceres-wheat model to various weather and non-weather parameters for wheat (cv.gw 496). The Journal of Agricultural Science, 1(2), 21–31.

Pohl, C. & van Genderen, J. (1998). Multisensor image fusion in remote sensing: concepts, methods and applications. Int. J. Remote Sensing, 19(5), 823–854.

Prevot, L., Dechambre, M., Taconet, O., Vidal-Madjar, D., Normand, M., & Galle, S. (1993). Estimating the characteristics of vegetation canopies with the airborne radar measurements. Int. J. Remote Sensing, 14, 2803–2818.

Prévot, L., Chauki, H., Troufleau, D., Weiss, M., Baret, F., & Brisson, N. (2003). Assimilating optical and radar data into the stics model for wheat. Agronomie, 23, 297–303.

Richards, J. (1993). Remote Sensing Digital Image Analysis - An Introduction. Springer-Verlag.

Richards, J. (2005). Analysis of remotely sensed data: The formative decades and the future. IEEE Trans. Geosci. Remote Sensing, 43(3), 422–432.

Richter, R. (1990). A fast atmospheric correction algorithm applied to landsat tm. Int. J. Remote Sensing, 11(1), 159–166.

Rickman, R. & Klepper, B. (1991). Environmentally driven cereal crop growth models. Annu. Rev. Phytopathol., 29, 361–380.

Riedel, T. & Schmullius, C. (2003). Impact of interception on the thematic analyses of sar data in agricultural areas. In International Geoscience and Remote Sensing Symposia, IGARSS.

Rosenqvist, A., Milne, A., & Zimmermann, R. (2003). Systematic data acquisitions-a prerequisite for meaningful biophysical parameter retrieval? IEEE Trans. Geosci. Remote Sensing, 41(7), 1709–1711.

Rosenzweig, C., Allen, L., Jones, J., Tsuji, G., & Hildebrand, P. (1995). Climate Change and Agriculture: Analysis of Potential International impacts. Amer. Soc. Agron.

Rouse, J., Haas, R., Schell, J., & Deering, D. (1978). Monitoring the vernal advancement and retrogradations (green wave effect) of natural vegetation. Technical report, Prog. Rep. RSC 1978-1, Remote Sensing Center, Texas A&M Univ., College Station, 93p. (NTIS No. E73-106393).

Saich, P., Cordey, R., Quegan, S., Williams, M., Baker, J., Luckman, A., Wielogorska, A., & Wooding, M. (1995). Sar retrieval algorithms for land applications. Technical report, ESA Contrac Report 10644/93/NL/NB.

Schiefer, S., Hostert, P., & Damm, A. (2006). Correcting brightness gradients in hyperspectral data from urban areas. Remote Sens. Environ., 1(1), 25–37.

Schmullius, C. & Nithack, J. (1995). Crop monitoring with multi-temporal airborne dlr e-sar images. In International Geoscience and Remote Sensing Symposia, IGARSS.

Schmullius, C., Nithack, J., & Kern, M. (1993). Comparison of multi-temporal ers-1 and airborne dlr e-sar image data for crop monitoring. In Proc. 2nd ERS-1 Symp. Hamburg, Germany.

Schoenwiese, C.-D. (1992). Praktische Statistik für Meterologen und Geowissenschaflter. Gebrüder Bornträger.

Schowengerdt, R. A. (1997). Remote Sensing - MODELS AND METHODS FOR IMAGE PREPROCESSING. Academic Press.

Siqueira, P. & Sarabandi, K. (1996). A numerically derived electromagnetic scattering model for grass grain heads. IEEE Trans. Geosci. Remote Sensing, 2(2), 1337–1339.

Spoenemann, J. & Schieche, B. (1997). Fernerkundung mittels satelliten als datenquelle der agrarstatistik am beispiel des landkreises göttingen. In Sonderdruck: Geographie in der Grundlagenforschung und als Angewandte Wissenschaft. Göttinger Akzente. Göttinger Geographische Abhandlungen 100.

Steven, M., Malthus, T., Baret, F., Xu, H., & Chopping, M. (2003). Intercalibration of vegetation indices from different sensor systems. Remote Sens. Environ., 88(4), 412–422.

Teilet, P. M., Barker, J. L., Markham, B. L., Irish, R. R.and Fedosejevs, G., & Storey, J. (2001). Radiometric cross-calibration of the landsat-7 etm+ and landsat-5 tm sensors based on tandem data set. Remote Sens. Environ., 78(1-2), 39–54.

Tewolde, H., Sistani, K., Rowe, D., Adeli, A., & Tsegaye, T. (2005). Estimating cotton leaf area index nondestructively with a light sensor. Agron. J., 97, 1158–1163.

Tilman, D., Cassman, K., Matson, P., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and intensive production practices. Nature, 418, 671–677.

Tilman, D., Fargione, J., Wolff, B., Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W., Simberloff, D., & Swackhamer, D. (2001). Forecasting agriculturally driven global environmental change. Science, 292, 281–284.

Trishchenko, A., Cihlar, J., & Li, Z. (2002). Effects of spectral response function on surface reflectance and ndvi measured with moderate resolution satellite sensors. Remote Sens. Environ., 81(1), 1–18.

Ulaby, F., Kouyate, F., & Fung, A. (1981). Microwave Remote Sensing: Active and Passive, Vol. I Microwave Remote Sensing Fundamentals and Radiometry. Addison-Wesley Publishing Co., Reading, MA.

Ulaby, F., Moore, R., & Fung, A. (1982). Microwave Remote Sensing. Active and Passive. Radar Remote Sensing and Surface Scattering and Emission Theory. Artech House, Inc.

Ulaby, F., Moore, R., & Fung, A. (1986). Mircowave Remote Sensing. Active and Passive. From Theory to applications. Artech House, Inc.

Ulaby, F., Sarabandi, K., McDonald, K., Whitt, M., & Dobson, M. (1990). Michigan microwave canopy scattering model. Int. J. Remote Sensing, 11(7), 1223–1253.

Ulaby, F. & Wilson, E. (1985). Microwave attenuation properties of vegetation canopies. IEEE Trans. Geosci. Remote Sensing, GE-23(5), 746–753.

Ustin, S. (2004). Remote Sensing for Natural Resource Management and Environmental Monitoring. John Wiley & Sons.

van Leeuwen, W., Orr, B., Marsh, S., & Herrmann, S. (2006). Multi-sensor ndvi data continuity: Uncertainties and implications for vegetation monitoring applications. Remote Sens. Environ., 100(1), 67–81.

Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: the sail model. Remote Sens. Environ., 16(2), 125–141.

von Thünen, J. (1910). Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Fischer.

Waggoner, P. (1995). How much land can ten billion people spare for nature? Technol. Soc., 17, 17–34.

Waring, R., Way, J., Hunt, E., Morrissey, L., Ranson, J., Weishampel, J., Oren, M., & Franklin, S. (1995). Imaging radar for ecosystem studies. Bioscience, 45, 715–723.

Waske, B. & Benediktsson, J. (2007). Fusion of support vector machines for classification of multisensor data. IEEE Trans. Geosci. Remote Sensing, 45(12), 1–9.

Watson, D. (1947). Comparative physiological studies in the growth of field crops. i. variation in net assimilation rate and leaf area between species and varieties, and within and between years. Annuals of Botany, 11, 41–76.

Wiegand, C., Gerbermann, A., Gallo, K., Blad, B., & Dusek, D. (1990). Multi-site analyses of spectral-biophysical data for corn. Remote Sens. Environ., 33(1), 1–16.

Wiegand, C., Maas, S., Aase, J., Hatfield, J., Pinter, P., Jackson, R., Kanemasu, E., & Lapitan, R. (1992). Multi-site analyses of spectral-biophysical data for wheat. Remote Sens. Environ., 42(1), 1–21.

Wit, C. d. (1965). Photosynthesis of leaf canopies. Technical report, Agricultural Research Report 663, Pudoc, Wageningen, the Netherlands.

Witzenberger, A., Hack, H., & van den Boom, T. (41). Erläuterungen zum bbch-dezimal-code für die entwicklungsstadien des getreides -mit abbildungen. Gesunde Pflanzen, 1989, 384–388.

Wood, D., McNairn, H., Brown, R., & Dixon, R. (2002). The effect of dew on the use of radarsat-1 for crop monitoring , choosing between ascending and descending orbits. Remote Sens. Environ., 80(2), 241–247.

Wood, S., Sebastian, K., & Scherr, S. (2000). Pilot analysis of global ecosystems: Agroecosystems. Technical report, International Food Policy Research Institute and World Resources Institute.

Woodhouse, I. H. (2006). Introduction to Microwave Remote Sensing. Taylor & Francis.

Wooding, M., Attema, E., Aschbacher, J., Borgeaud, M., Cordey, R., de Groof, H., Harms, J., Lichtenegger, J., Nieuwenhuis, G., Schmullius, C., & Zmuda, A. (1995). Satellite radar in agriculture. experience with ers-1. Technical report, ESA SP-1185, ISBN 90-9092- 339-3.

Worldbank (2000). Agriculture and achieving the millenium development goals. Technical report, World Bank/International Food Policy Research Institute. Report 32729-GLB, Washington, USA.

A ATTACHMENT

Table A.1: Translation coefficients for sixth order polynomial NDVI intercalibration. Where the subscripts SOURCE and TARGET indicate the sensor to be translated and the reference sensor.

SOURCE NDVI	TARGET NDVI SPOT 5	Landsat 5 TM	QuickBird	ASTER
SPOT 5		a_0=0.005	a_0=-0.001	a_0=0.001
		a_1=0.833	a_1=0.867	a_1=0.832
		a_2=0.278	a_2=0.109	a_2=0.153
		a_3=0.247	a_3=0.157	a_3=0.293
		a_4=-0.858	a_4=-0.292	a_4=-0.455
		a_5=0.381	a_5=0.042	a_5=-0.033
		a_6=0.082	a_6=0.068	a_6=0.168
Landsat 5 TM 5	a_0=-0.006		a_0=-0.006	a_0=-0.005
	a_1 =1.181		a_1=1.033	a_1=1.001
	a_2=-0.282		a_2=-0.154	a_2=-0.114
	a_3=-0.392		a_3=-0.159	a_3=-0.012
	a_4=1.009		a_4=0.609	a_4=0.448
	a_5=-0.178		a_5=-0.211	a_5=-0.334
	a_6=-0.313		a_6=-0.137	a_6=0.001
QuickBird	a_0=0.001	a_0=0.006		a_0=0.001
	a_1 =1.144	a_1=0.969		a_1=0.966
	a_2=-0.139	a_2=0.186		a_2=0.052
	a_3=-0.187	a_3=0.088		a_3=0.144
	a_4=0.424	a_4=-0.721		a_4=-0.209
	a_5=-0.115	a_5=0.535		a_5=-0.062
	a_6=-0.081	a_6=-0.038		a_6=0.118
ASTER	a_0=0.001	a_0=0.005	a_0=-0.001	
	a_1=1.178	a_1=1.00	a_1=1.033	
	a_2=-0.194	a_2=0.122	a_2=-0.052	
	a_3=-0.331	a_3=-0.042	a_3=-0.136	
	a_4=0.635	a_4=-0.480	a_4=0.203	
	a_5=-0.052	a_5=0.528	a_5=-0.052	
	a_6=-0.202	a_6=-0.125	a_6=-0.109	

Table A.2: Fitted coefficients for the multiple regression analysis of fresh biomass using SAR data.

all input parameters					
A	B	C	D	E	F
ERS-2 (VV)					
-5.58	-0.35	8.66	0.40	0.26	-0.30
ASAR (HH)					
5.43	-8.65	-3.06	0.38	-0.20	-3.51
ASAR (HV)					
4.79	-39.51	-2.19	0.41	-0.09	-3.96
ASAR (HH/HV)					
3.18	3.44	-1.77	0.38	-0.12	-3.95
without roughness					
ERS-2 (VV)					
-5.31	-0.42	10.61		0.24	
ASAR (HH)					
5.81	4.38	-3.21		-0.37	
ASAR (HV)					
8.02	-48.06	-3.53		-0.47	
ASAR (HH/HV)					
8.34	-8.74	-4.48		-0.44	

Table A.3: Fitted coefficients for the multiple regression analysis of plant water content using SAR data.

all input parameters					
A	B	C	D	E	F
ERS-2 (VV)					
-8.10	-0.39	13.61	0.15	0.23	0.42
ASAR (HH)					
2.13	-7.14	-0.63	0.39	0.18	-1.97
ASAR (HV)					
1.85	-36.45	0.07	0.41	0.13	-2.42
ASAR (HH/HV)					
-0.08	5.23	-0.02	0.40	0.24	-2.07
without roughness					
ERS-2 (VV)					
-7.36	-0.41	12.85		0.24	
ASAR (HH)					
2.57	6.17	-1.46		-0.01	
ASAR (HV)					
5.12	-45.23	-1.49		-0.13	
ASAR (HH/HV)					
5.32	-8.98	-2.26		-0.10	

Table A.4: Fitted coefficients for the multiple regression analysis of dry biomass using SAR data.

all input parameters					
A	B	C	D	E	F
ERS-2 (VV)					
2.48	0.03	-4.84	0.26	0.03	-0.74
ASAR (HH)					
3.29	-1.52	-2.41	-0.02	-0.36	-1.52
ASAR (HV)					
3.11	-3.74	-2.16	-0.02	-0.35	-1.64
ASAR (HH/HV)					
3.24	-1.81	-1.75	0.01	-0.36	-1.86
without roughness					
ERS-2 (VV)					
2.02	-0.01	-2.15		-0.01	
ASAR (HH)					
3.25	-1.80	-1.76		-0.37	
ASAR (HV)					
2.93	-3.27	-2.07		-0.34	
ASAR (HH/HV)					
3.05	-1.21	-2.25		-0.34	

Table A.5: Fitted coefficients for the multiple regression analysis of LAI using SAR data.

all input parameters					
A	B	C	D	E	F
ERS-2 (VV)					
-4.09	-0.13	9.66	0.07	0.02	-0.09
ASAR (HH)					
1.92	1.70	-0.62	0.04	-0.07	-1.77
ASAR (HV)					
2.23	1.22	-0.94	0.02	-0.08	-1.62
ASAR (HH/HV)					
3.06	-5.47	0.01	0.07	-0.13	-2.22
without roughness					
ERS-2 (VV)					
-3.95	-0.14	9.50		0.03	
ASAR (HH)					
3.54	-5.29	-0.25		-0.17	
ASAR (HV)					
2.33	0.22	-0.90		-0.08	
ASAR (HH/HV)					
2.09	1.65	-0.68		-0.08	

Table A.6: Fitted coefficients for the Water Cloud model (WCM) and the retrieval of fresh biomass (with (+VEGE) and without vegetation roughness considerations).

WCM+VEGE			
A	B	C	D
ERS-2 (VV)			
0.36	0.63	-16.55	0.613
ASAR (HH)			
-0.16	0.47	-12.72	-0.60
ASAR (HV)			
-0.35	0.52	-19.97	-1.75
ASAR (HH/HV)			
-0.24	0.05	-7.24	-1.15
WCM			
ERS-2 (VV)			
	0.86	-21.89	0.56
ASAR (HH)			
	0.49	-12.96	-0.58
ASAR (HV)			
	0.57	-20.48	-1.64
ASAR (HH/HV)			
	0.08	-7.79	-1.10

Table A.7: Fitted coefficients for the Water Cloud model (WCM) and the retrieval of plant water content (with (+VEGE) and without vegetation roughness considerations).

WCM+VEGE			
A	B	C	D
ERS-2 (VV)			
-0.49	0.54	-14.83	0.54
ASAR (HH)			
-0.43	0.56	-12.61	-0.55
ASAR (HV)			
-1.91	0.26	-18.38	-2.05
ASAR (HH/HV)			
-1.04	-0.22	-6.96	-1.37
WCM			
ERS-2 (VV)			
	0.66	-18.15	0.68
ASAR (HH)			
	0.64	-12.56	-0.44
ASAR (HV)			
	0.64	-22.53	-1.56
ASAR (HH/HV)			
	-0.01	-9.00	-1.12

Table A.8: Fitted coefficients for the Water Cloud model (WCM) and the retrieval of dry biomass (with (+VEGE) and without vegetation roughness considerations).

WCM+VEGE			
A	B	C	D
ERS-2 (VV)			
-1.62	0.40	-17.74	0.46
ASAR (HH)			
-1.20	0.88	-11.18	-1.03
ASAR (HV)			
-0.50	1.79	-19.14	-2.17
ASAR (HH/HV)			
1.01	0.95	-9.26	-1.08
WCM			
ERS-2 (VV)			
	0.86	-21.89	0.56
ASAR (HH)			
	1.25	-14.56	-0.86
ASAR (HV)			
	1.95	-20.53	-2.10
ASAR (HH/HV)			
	0.70	-5.98	-1.23

Table A.9: Fitted coefficients for the Water Cloud model (WCM) and the retrieval of LAI (with (+VEGE) and without vegetation roughness considerations).

WCM+VEGE			
A	B	C	D
ERS-2 (VV)			
-0.95	0.48	-13.06	0.32
ASAR (HH)			
-1.86	0.39	-10.18	-0.91
ASAR (HV)			
-1.52	1.12	-17.01	-1.79
ASAR (HH/HV)			
0.31	0.66	-6.91	-0.85
WCM			
ERS-2 (VV)			
	0.83	-20.07	0.56
ASAR (HH)			
	0.93	-15.26	-0.44
ASAR (HV)			
	1.55	-21.08	-1.41
ASAR (HH/HV)			
	0.62	-5.86	-0.97

Table A.10: Fitted coefficients for the combined (multi-sensoral) biophysical parameter retrieval, using the different polarizations and the NDVI.

VV			
	A	B	C
FM	-6.43	-0.57	4.60
DM	-0.82	-0.19	-0.46
LAI	9.43	-0.35	-12.92
PWC	-5.61	-0.38	5.06
HH			
	A	B	C
FM	2.47	-0.46	-3.92
DM	3.34	-0.14	-4.93
LAI	-2.71	0.03	5.53
PWC	-0.87	-0.32	1.01
HV			
	A	B	C
FM	10.45	-0.46	-3.92
DM	6.25	0.26	-0.83
LAI	-3.32	-0.08	4.12
PWC	4.20	-0.77	.2.31
ratio			
	A	B	C
FM	1.90	.0.58	7.69
DM	2.79	0.20	-0.71
LAI	-2.55	-0.04	4.65
PWC	-0.89	0.37	8.40

Table A.11: Fitted coefficients for the combined (multi-sensoral) biophysical parameter retrieval, using the VV polarization regression, the CLAIR model and respective weighting factors.

early phenological stages						
	α	$WDVI_\infty$	A	B	C	E
FM	1000	1000	-1.23	-20.10	17.02	0.48
DM	900	950	-1.94	-7.70	10.01	0.15
LAI	980	850	1.39	0.01	2.14	0.07
PWC	1100	1050	-0.61	-26.35	21.47	0.76
late phenological stages						
	α	$WDVI_\infty$	A	B	C	E
FM	1029	1044	-212.23	5.19	544.38	5.31
DM	999	978	118.58	-43.51	-273.73	-4.60
LAI	1080	1090	-155.01	-2.49	389.29	3.42
PWC	1046	4052	-358.06	32.42	898.65	14.47

Table A.12: Fitted coefficients for the combined (multi-sensoral) biophysical parameter retrieval, using the Water Cloud Model (VV polarization), the CLAIR model and respective weighting factors.

early phenological stages					
	α	$WDVI_\infty$	B	C	D
FM	1011	1030	1.14	-8.62	1.47
DM	4482	4571	0.01	-7.39	1.47
LAI	1090	1010	0.84	3.01	0.27
PWC	1090	1110	1.61	-8.41	1.47
late phenological stages					
	α	$WDVI_\infty$	B	C	D
FM	901	1013	0.67	-7.21	2.19
DM	1032	1035	2.74	46311	-3645
LAI	1000	1005	0.28	-14.23	0.75
PWC	1001	999	0.76	-17.72	3.10