ASSESSING THE QUALITY OF TEA BY HYPERSPECTRAL TECHNIQUES

Meng Bian

Examining committee:

Prof.dr. A. Veldkamp	University of Twente
Prof.dr.ing. W. Verhoef	University of Twente
Prof.dr. E. Milton	University of Southampton
Dr.ir. J.G.P.W. Clevers	Wageningen University

ITC dissertation number 230
ITC, P.O. Box 6, 7500 AA Enschede, The Netherlands

ISBN 978-90-6164-358-6
Cover designed by Benno Masselink
Printed by ITC Printing Department

UNIVERSITY OF TWENTE.

FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

ASSESSING THE QUALITY OF TEA BY HYPERSPECTRAL TECHNIQUES

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday July 10, 2013 at 12:45 hrs

by

Meng Bian

born on 11 December 1981

in Shandong, China

This thesis is approved by

Prof. dr. Andrew K. Skidmore,	Promotor
Dr. Tiejun Wang,	Assistant promotor

ACKNOWLEDGEMENTS

My heartiest thanks go to Prof. Andrew K. Skidmore, my Ph.D. promotor and supervisor in The Netherlands, for his strict training and demands in scientific contents, incisive advice and patient guidance in my research. He devoted a lot of time to read all the drafts seriously, and made prompt and constructive comments. More important, he has greatly influenced me to think the sense of science. Without his foresight and continuous support, the thesis would not have come to fruition. Everything he taught me I will cherish for a lifetime.

My tremendous gratitude goes to Prof. Yaolin Liu and Prof. Yanfang Liu from Wuhan University, my supervisors in China, for their guidance and encouragements in all respects of the study. Their advice and support helped me in the evolution of my research.

I am so grateful to Dr. Jan de Leeuw, Dr. Martin Schlerf and Dr. Tiejun Wang of ITC, the excellent supervising team behind me, for their generous assistance, invaluable suggestions and constant encouragement throughout my study. Many thanks are due to their kind support and pertinent suggestions on the drafts of the proposal and the thesis. I am so lucky to have their supervisions and I deeply appreciate their unselfishness support and guidance.

Sincere gratitude is extended to all the teachers and students in The Netherlands and China who took part in the study for their generous help and important contributions to this work, in particular Prof. Pingxiang Li, Prof. Liangpei Zhang, Prof. Tianming Yang, Prof. Dejiang Ni, Ms. Rong Zeng, Dr. Yi Cen, Dr. Tao Chen and Dr. Yuancheng Huang. Without their participation and kind support, this study would not have been completed.

I would like to thank International Institute for Geo-information Science and Earth Observation (ITC), University of Twente, The Netherlands, and School of Resource and Environmental Science, Wuhan University, China, for providing me with the opportunity to conduct my research and develop my skills. I also deeply appreciate the financial support by the two organizations.

My sincere thanks also go to many teachers and friends who shared my ups and downs at ITC and Wuhan University: many thanks to Prof. Guofeng Wu and his wife Wenxiu Gao, for the great help and support from them; I thank Dr. David Rossiter, my teacher and friend, for giving us impressive class on scientific writing and for his great help and clear explanation on statistics; I thank Dr. Michael Weir and Prof. Bob Su, for their always care and support, particularly when I have difficulties in the research; Ms. Esther Hondebrink, Ms. Loes Colenbrander, Ms. Laura Dente, Dr. Rishiraj Dutta, Ms. Lichun Wang, Ms. Linlin Pei, and other staff and friends, thanks a lot for giving me generous help and tremendous support.

Special thanks go to my colleagues in NRS, in particular Dr. Roshanak Darvishzadeh, Ms. Sabrina Carvalho, Mr. Anas Fauzi, Mr. Atkilt Girma, Mr. Babak Naimi, Mr. Aidin Niamir, Ms. Claudia Pittiglio, Ms. Maitreyi Sur, Mr. Ben Wielstra, Ms. Ha Nguyen, Dr. Tyas Basuki, Dr. Moses Cho, Dr. Jelle Ferwerda, Dr. Mhosisi Mosacha, Dr. Nicky Knox, Ms. Filiz Bektas Balcik, Mr. Ajay Bhakta Matherma, and to the warm Chinese group at ITC. They discussed with me a lot on my research and gave me lots of invaluable suggestions and helps in both of the study and life. I got so much happiness and encourages from all of them during my stay in The Netherlands. Their priceless friendship and assistance will always be treasured in my heart.

My deep gratitude goes to Ms. Eva Skidmore, who did excellent English editing to help me publish my first paper. I also appreciate her generous and patient help when I had problems on English writing.

Finally, I am much obliged to my dear parents for their great support throughout my life. I would also like to express my heartfelt appreciation to my husband, Teng Fei for his company, tolerance and great help during this learning experience. The deep love and practical support I enjoy from my families have always given me faith and strength. I only want to say: with all my heart, I love you!

Contents

Acknowledgements ..... **i**

Chapter 1 General introduction **1**
1.1 Background and problem statement: tea quality evaluation 2
1.2 Biochemical components as indicators of tea quality 6
1.3 Hyperspectral remote sensing of vegetation biochemical characteristics 8
1.4 Developments in techniques for estimating biochemicals using hyperspectral data 10
1.5 Research objectives and questions 12
1.6 Study site and tea cultivars 13
1.6.1 Study area *13*
1.6.2 Tea cultivars *14*
1.7 Outline of thesis 15

Chapter 2 Reflectance spectroscopy of biochemical components as indicators of tea quality **17**
2.1 Introduction 19
2.2 Material and methods 21
2.2.1 Study area *21*
2.2.2 Sampling design *21*
2.2.3 Spectral measurements *22*
2.2.4 Biochemical assay *23*
2.2.5 Data analysis *23*
2.3 Results 25
2.4 Discussion 30
2.5 Conclusions 33

Chapter 3 Estimation of tea quality from canopy hyperspectral measurements **35**
3.1 Introduction 37
3.2 Materials and methods 40
3.2.1 Plant material and growing conditions *40*
3.2.2 Canopy spectral measurements and preprocessing *41*
3.2.3 Chemical analysis *41*
3.2.4 Data analysis *42*
3.3 Results 46
3.3.1 Descriptive statistics of foliar chemistry *46*
3.3.2 Training and testing the neural networks based on the wavebands selected by the successive projections algorithm (SPA-ANN) *47*
3.3.3 Comparison of models with SPA, PCA and random selections *49*

3.4 Discussion and conclusions 51

Chapter 4 Predicting foliar biochemistry of tea using hyperspectral data from powder, leaf to canopy 55

4.1 Introduction 57
4.2 Methods and materials 59
4.2.1 Study area 59
4.2.2 Sampling design 60
4.2.3 Data collection 60
4.2.4 Data analysis 64
4.3 Results 66
4.4 Discussion 72
4.5 Conclusions 74

Chapter 5 Tea-PROSPECT: a leaf optical model for tea incorporating polyphenols 77

5.1 Introduction 79
5.2 Materials and Methods 81
5.2.1 Study area and sampling strategies 81
5.2.2 Data collection 82
5.2.3 Improved PROSPECT model for tea (tea-PROSPECT) 84
5.2.4 Spectral simulation and model inversion 86
5.3 Results 88
5.4 Discussion 93
5.4.1 Leaf spectral simulation of tea 93
5.4.2 Model inversion for the retrieval of total tea polyphenols 94
5.5 Conclusions 96

Chapter 6 Synthesis - assessing the quality of tea from laboratory NIRS to field spectral measurements 97

6.1 Introduction 99
6.2 Laboratory level 102
6.3 Field level 104
6.4 From laboratory to field 105
6.5 Leaf optical model for tea species 108
6.6 Practical relevance of this research 110
6.6.1 For tea plantations managers (individual plantations) 111
6.6.2 For managers of tea plantations and processing plants 111
6.6.3 For Ministry of Agriculture (Domestic tea production and international tea trade) 111
6.6.4 Contributions to agricultural and food studies 112
6.7 General conclusion 112
6.8 Future perspectives 113

Bibliography **115**
Author's Biography **137**
Summary **139**
Samenvatting **141**
ITC Dissertation List **143**

CHAPTER 1

GENERAL INTRODUCTION

1.1 Background and problem statement: tea quality evaluation

Tea is made from young shoots and leaves of the plant *Camellia sinensis* (Willson and Clifford 1992). This plant is native to East and South Asia. Depending on the techniques used in the production and processing, tea can be divided into four basic types, namely green tea, black tea, oolong tea, and white tea (Liu and Yue 2005). In the early 17th century, green tea was first introduced into Europe through Amsterdam. During the following century, tea became fashionable and was introduced in France, Germany, England, America, Russia, as well as other countries (Martin 2007). With demand increasing, cultivation was extended into other tropical regions (Acland 1971; Dutta et al. 2010; Willson and Clifford 1992). Tea consumption has risen further in recent years, and has become the most popular beverage consumed worldwide besides water, as a result of its special flavour and possible beneficial health effects (Mukhtar and Ahmad 2000; Sin 2007).

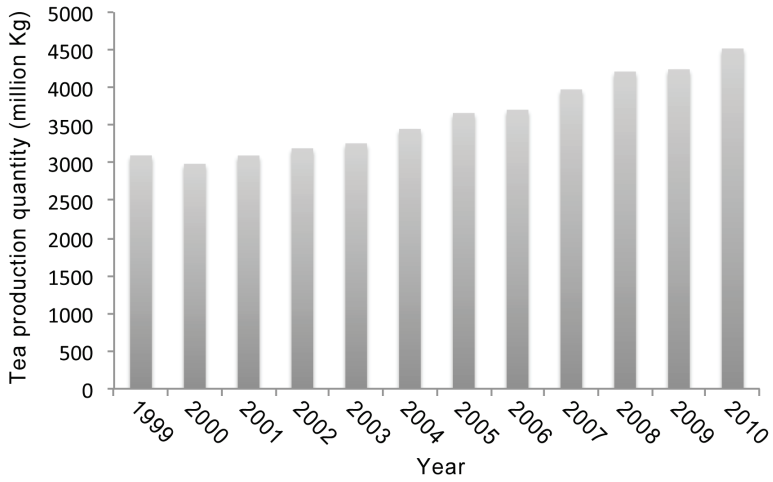

Figure 0-1 World tea production (unit: million kilograms) from 1999 to 2010. Data source: (FAO 2012)

Figure 0-1 shows the quantity of tea produced from 1999 to 2010, based on public data of the Food and Agriculture Organization of the United Nations. In 2006, about 3,700 million kilograms of tea have been produced all over the world, and by 2010, this number has increased to more than 4,500 million kilograms. Tea is an important cash crop in many countries, especially in

developing countries. The top three countries producing tea are China, India, and Sri Lanka (Figure 0-2), according to 2009 estimates by the Food and Agriculture Organization of the United Nations (2012) and the 2009 tea statistics by the Tea Board of India. The tea industry creates considerable economic benefits, employing a large labour force; for example, the Tea Board of India statistics show that at least one million labourers were employed in the tea industry in 2007 (Dutta 2011). As a tea garden is a complex ecosystem of its own, as well as a sub-system of larger ecosystems (Willson and Clifford 1992), good management of tea plantations not only maintains the ecological balance within the tea garden, but also contributes to a healthier broader ecosystem. In 2010, more than 3 million hectares (area harvested) of land were under tea cultivation globally, and the gross production value of tea (world total) increased to more than 9,000 million US$ (FAO 2012).

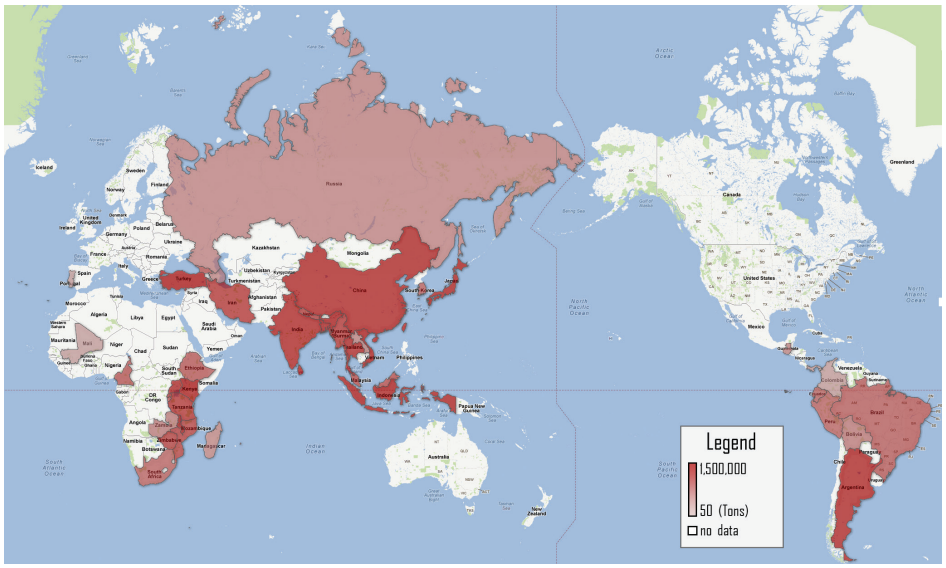

Figure 0-2 Tea production map (by country) in 2009. China, India and Sri Lanka are ranked top three tea producing countries. Data source: (FAO 2012).

With the increase in consumption, quality control of tea is becoming more and more important (Alaerts et al. 2012 In press; Chandini et al. 2011; Horie and Kohata 1998; Singhal et al. 1997). More attention has been paid to not only the quantity but also the quality of tea, as the quality is an essential factor influencing the market value of tea (Yan 2007). Many national and international

authorities are developing criteria for quality factors of tea (Chen et al. 2006; Thanh Mai et al. 2012). Meanwhile, a series of international standards have been set to establish scientific and uniform methods to determine characteristic substances of green and black tea (such as ISO 10727:2002, ISO 14502-1:2005 and ISO 14502-2:2005), or define a minimum biochemical content to indicate tea quality (such as ISO 3720 and ISO 11287).

Tea's quality is mainly reflected in its smell and taste, which are generated by volatile and non-volatile organic compounds present in the tea (Chen et al. 2008b; Dutta et al. 2003; Liang et al. 2003). The literature indicates that the quality of tea can be influenced at two different stages: during the growing period of the tea plant and at the tea processing stage (Obanda et al. 1997; Ravichandran and Parthiban 2000; Willson and Clifford 1992). After careful picking and selecting, the fresh tea leaves are dried, rolled, curled, and twisted for the production of green tea, or withered, curled, fermented, and dried for the production of black tea. Then, the quality of the tea is appraised and graded by professional tea tasters. The tea processing can be monitored by industrial quality control procedures. However, the biochemical properties of the fresh tea leaves in the growing stage are more difficult to control, though equally important (Wright et al. 2000). The concentration of foliar chemical compounds has an important impact on tea infusion's flavour, smell, and other factors that make up the tea's quality (Chen et al. 2006; Teranishi et al. 1999; Yamamoto et al. 1997).

Monitoring the chemical dynamics of fresh tea leaves provides an opportunity to predict the quality of the tea product before the leaves are being plucked (Bian et al. 2010). It will help to minimize subjectivity by tea judges and improve tea garden management, thus assuring consumers that the tea is of expected quality and strengthening confidence in the international tea trade. The current management from tea cultivation to tea harvesting largely depends on human experience (Acland 1971; Shankar et al. 2003). In practice, usually the young tender buds and leaves are plucked for producing a high-quality tea product. Compared with the older leaves, this part of the tea plant contains many more polyphenols, and a higher concentration of amino acids, forming the special taste of the tea beverage (Mitscher and Dolby 1997). However, information on the spatial distribution of tea quality is lacking. Although some

studies investigated the biochemical parameters of fresh tea leaves through sample surveys and traditional laboratory work (Gulati et al. 1999; Wei et al. 2003), these methods are difficult to implement in a fast, non-destructive way or over a large area. Therefore, a novel, cheaper, and effective method is needed to quantitatively monitor and estimate the biochemical parameters correlated to the quality of tea over large areas and in real time, for better management of tea plantations. Figure 0-3 demonstrates the scope of the study and points out the main aim.

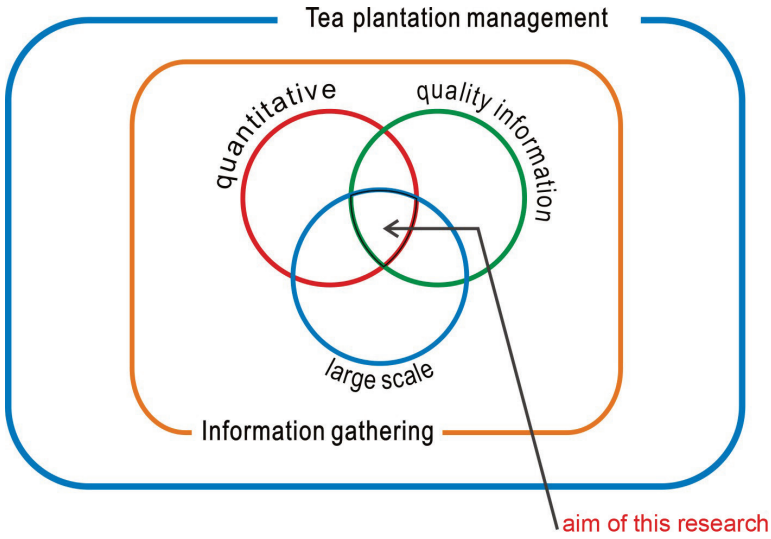

Figure 0-3 Diagrammatic summary of the scope of the study. The main objective is indicated by the arrow.

Remote sensing allows large and continuous radiometric measurements from which the biophysical and biochemical characteristics of vegetation may be derived. For tea species, several studies have been carried out to detect the biomass or predict the yield by geographic information system and remote sensing techniques (Rajapakse 2002; Rao et al. 2007), but few studies have worked on quality assessment of tea plants. If we want to estimate biochemical concentrations of tea with satisfactory accuracy using spectroscopy, traditional broadband remote sensing has its limitations, for it uses average spectral information across broadband widths resulting in loss of critical information available in specific narrow bands. To overcome these disadvantages, high-quality field and airborne sensors that can acquire hyperspectral data have been

developed and used for retrieving some biochemical parameters of vegetation from the reflected spectra (Ferwerda et al. 2005; Huber et al. 2008; Stagakis et al. 2010). The question posed is, can hyperspectral techniques be applied to assess the quality of tea and help explain tea quality variation in a spatial context?

1.2 Biochemical components as indicators of tea quality

Determining tea quality is traditionally the task of tea experts. An experienced tea taster looks at the colour of the tea infusion, smells the dry tea leaves and the liquid, then takes about 5 millilitre (a teaspoon) of tea, keeping the liquid in his or her mouth for a moment to allow the fragrance, taste, and other sensations to develop (Methodology of sensory evaluation of tea, GB/T23776-2009 of China). Finally, the flavour characteristics are graded by the tea taster. The leaf colour, size, and shape are also graded and, together with the flavour score, they quantify the overall quality. The whole procedure is called tea sensory evaluation (Stone and Sidel 2004; Togari et al. 1995). These quality aspects, especially the taste and smell, are largely determined by the biochemical components in fresh tea leaves, the raw material for tea (Nakagawa 1970; Owuor and Obanda 2007; Taylor et al. 1992; Wei et al. 2011). In the tea industry, research has been carried out to investigate the effect of different foliar biochemicals on tea quality (Hussain et al. 2011; Liang and Liu 2008; Obanda et al. 1997; Wang and Ho 2009).

Fresh tea leaves contain caffeine, tea polyphenols, tea polysaccharides, and necessary nutrients, such as protein, amino acids, lipids, and vitamins. Generally, four chemical components – free amino acids, total tea polyphenols, soluble sugars, and caffeine – are considered indicators of tea quality (Ruan et al. 2010; Yamamoto et al. 1997; Yuan 2003). Unlike other natural vegetation, tea plants have been cultivated for a long time. Most propagation attempts were restricted to the selection of elite mother bushes or progenitors from within the natural hybrid population of tea (Satyanarayana and Sharma 1986). The selection process has been limited to those plants (clones) which can produce an acceptable cup of tea, thus shaping the chemical characteristics of today's tea (Agarwal et al. 1992; Banerjee 1988). The special flavour and astringency of a tea brew is the result of the total tea polyphenols, which amount to 18-36% of

the weight of the total dry matter of a tea leaf. Compared with most other plants, polyphenols have a high concentration in tea leaves (Graham 1992). They easily oxidize and may react with some other chemical constituents during the processing, thereby influencing the quality of brewed tea (Obanda et al. 2001). Over three-quarters of tea polyphenols are flavanols. The main flavanols in tea leaves are catechins, including epicatechin (EC), epicatechin gallate (ECG), epigallocatechin (EGC) and epigallocatechin gallate (EGCG) (Lin et al. 1996; Yang and Wang 1993). Also, total tea polyphenols may be of benefit to human health as antioxidants (Frei and Higdon 2003; McKay and Blumberg 2002; Sharangi 2009). The freshness and brothy taste of tea originates in the amino acid fraction, especially the unique amino acid in tea, theanine, which accounts for more than 50% of the total amino acids in tea leaves (Willson and Clifford 1992). Theanine belongs to the free amino acids, and has been proved to have a relaxing effect on human beings (Juneja et al. 1999; Kobayashi et al. 1998). The free amino acids have been regarded as having a positive effect on the taste of green tea (Hung et al. 2010; Nakagawa 1975). Soluble sugars add sweetness and viscosity to tea (Ding et al. 2002; Scharbert and Hofmann 2005). Caffeine, the most important alkaloid in tea, contributes to the bitterness of the brew through complexation with the flavanols, and may stimulate the nervous system of human beings (Dutta et al. 2011; Wörth et al. 2000). For some black teas, the high soluble caffeine (3.4-3.9%) causes their high briskness (Borse et al. 2002). Figure 0-4 presents the structural formulae of some of the major substances composing the main four biochemical components.

The concentrations of total tea polyphenols, free amino acids, soluble sugars, and caffeine vary according to the choice of tea clone, season, picking method and to the environmental conditions such as soil (Dang 2005; Hilton et al. 1973; Kumar and Rajapaksha 2005; Owuor and Obanda 1998). In this study, I focused on retrieving these four biochemicals, because of their importance to the taste and smell of tea, and therefore to the quality.

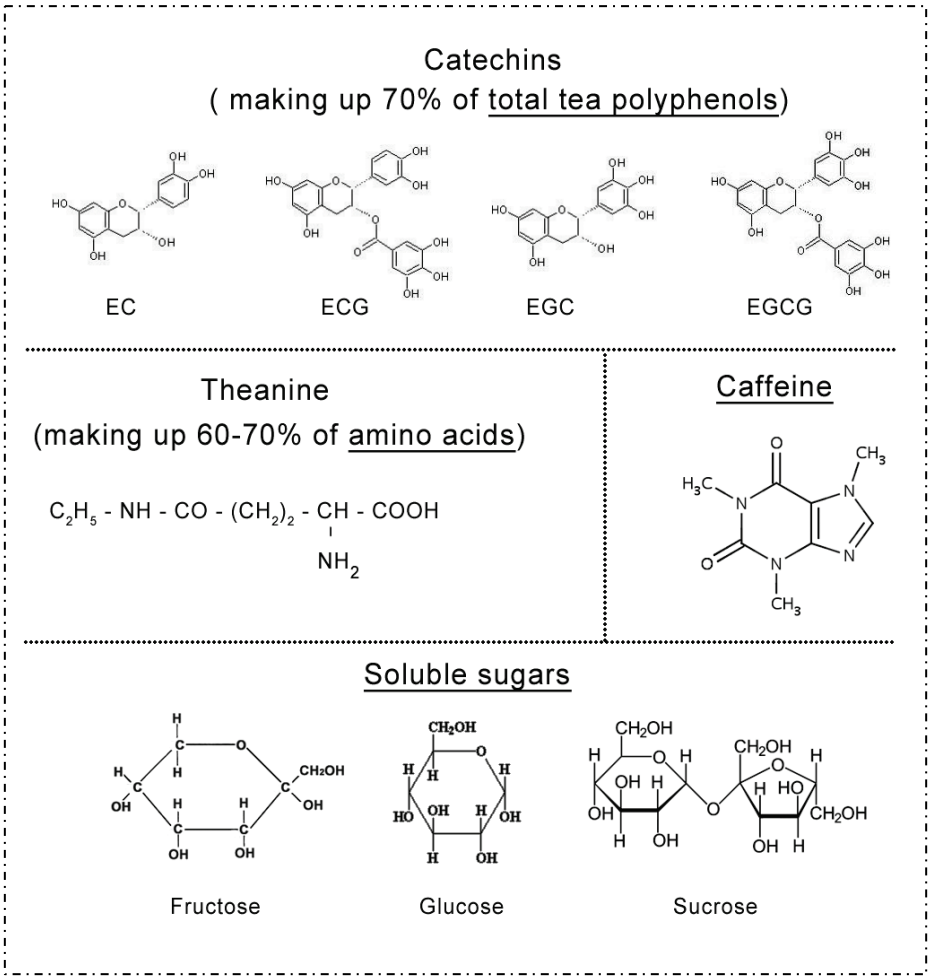

Figure 0-4 Structure of chemical compounds as quality indicators of tea

1.3 Hyperspectral remote sensing of vegetation biochemical characteristics

With the development of better sensors and faster computation techniques, a new field of research is emerging in remote sensing, called "hyperspectral remote sensing". This new earth observation technique emerged in the 1980s originates in a research plan of the "Imaging Spectrometer" project (Green et al. 1998). Hyperspectral reflectance consists of a large number of narrow spectral bands, usually over 100 bands along the visible, near infrared and shortwave infrared spectral plane (400-2500 nm). Most hyperspectral sensors have a spectral resolution of no more than 10 nm (Chang 2003; Green et al. 1998). The

narrow spectral channels in hyperspectral techniques make it possible to detect small changes in narrow absorption features that are caused by the biochemical and biophysical characteristics of vegetation (Broge and Leblanc 2001; Mutanga et al. 2009; Skidmore et al. 2010; Thenkabail et al. 2000).

The absorption features of vegetation are the result of electron transitions in chlorophyll and of the bending and stretching of the chemical bonds in water and other chemicals (Curran 1989). Molecular vibration of the C-H, N-H, O-H, C-N and C-C bonds within organic compounds cause strong absorption at wavelengths of 2500-15000 nm (middle-infrared). The overtones and harmonics result in absorption in the near infrared (1000-2500 nm), allowing organic compounds to be detected (Weyer 1985). Detailed spectroscopic measurements of dried and ground leaves were made in the 1960s and1970s by researchers at the U.S. Department of Agriculture (USDA) using laboratory near infrared spectroscopy (NIRS). About 42 absorption features from 400 to 2500 nm have been correlated with the concentration of organic compounds (e.g., cellulose, lignin, protein, oil, starch) in dried leaves (Curran 1989). Given the advent of hand-held, airborne and space-borne spectrometers with high spectral resolution, the procedure using NIRS has been gradually extended from the laboratory to the field, from dried leaf samples to living vegetation (Adam et al. 2010; Fava et al. 2009; Mutanga and Ismail 2010; Mutanga and Kumar 2007; Ramoelo et al. 2011a). Estimation of foliar biochemicals has improved significantly from early broad-band sensor attempts (Kokaly et al. 2009). To date, hyperspectral remote sensing techniques have been successfully applied to forest, forage and agricultural crops for retrieving parameters such as chlorophyll, nitrogen, cellulose and lignin, particularly at leaf level and with varying success at canopy level (Majeke et al. 2008). Most studies on hyperspectral remote sensing of biochemistry are based on leaf reflectance measured under laboratory conditions, because water and other canopy structure parameters, as well as environmental conditions complicate the detection of biochemical absorption features in living plants (Majeke et al. 2008).

1.5 Developments in techniques for estimating biochemicals using hyperspectral data

The task of retrieving biochemicals of vegetation from spectra is made difficult by the fact that reflectance, absorption, or transmittance at each wavelength is usually affected by more than one characteristic of the vegetation (Asner et al. 2000). Thus, various methodologies have been devised to obtain the maximum spectral signal from the biochemicals of interest, meanwhile minimizing the effect of spectral variability that is independent of the biochemical concentration of interest (Curran et al. 2001; Ferwerda et al. 2005; Mutanga et al. 2004b; Workman and Springsteen 1998). Most of the application methods have their foundations in near-infrared spectroscopy (NIRS) and the geological science domain.

One of the most common approaches to estimating vegetation parameters from remotely sensed data is the statistical approach, which continues to dominate hyperspectral remote sensing of biochemistry. Using a statistical approach, an empirical model is typically built by regressing the biochemical of interest with either the spectral reflectance at a specific wavelength, or a spectral index (in the case of a univariate regression) (Majeke et al. 2008). As a narrow band vegetation index derived from hyperspectral data, the normalized difference vegetation index (NDVI) is calculated using any possible combination of two wavebands in the spectral region, and has been applied in studies by Ferwerda et al. (2005) and Darvishzadeh et al. (2008c) to quantify the biochemical and biophysical characteristics of mopane trees and grass in savanna. Cho and Skidmore (2006) proposed a new technique for extracting the red-edge position from hyperspectral data. Based on this technique, Darvishzadeh et al. (2008c) successfully estimated leaf area index (LAI) and chlorophyll content for a heterogeneous grassland using hyperspectral measurements. Spectral vegetation indices and red edge position belong to univariate regression. Multivariate techniques involving stepwise multiple linear regression and partial least square regression have been addressed in some other studies, in which spectral information of part or all of the spectrum has been integrated to estimate vegetation biochemical properties (Carrascal et al. 2009; Dutta et al. 2011). Compared with the univariate regression method, a multivariate approach could exploit the richer information content of hyperspectral data. Moreover, to model complex nonlinearity that may exist between independent (waveband) and

dependent (biochemical) variables, Mutanga and Skidmore (2004b) integrated imaging spectroscopy and neural networks for the first time to map grass quality in South Africa. Artificial neural networks are artificial intelligence techniques, which are based on the perceived thought processes of the human brain (Atkinson and Tatnall 1997).It has been shown to be a powerful tool in modelling a variety of nonlinear behaviour, involving modelling a large range of transfer functions (Atkinson and Tatnall 1997; Islam and Kothari 2000; Mas and Flores 2008). But the large number of wavebands in hyperspectral remote sensing increases the challenge of retrieving biochemicals. In previous studies using hyperspectral remote sensing in biochemistry, artificial neural networks have not frequently featured (Mutanga and Skidmore 2004b).

Another approach to retrieving vegetation parameters is the physically based model (Liang 2004). It stems from our understanding of the physical mechanism of radiation transferring, which resulted in a better understanding of the interaction of light with plant leaves (Healey and Slater 1999; Jacquemoud et al. 2009; Sandmeier et al. 1998). Vegetation properties (i.e. albedo, mesophyll structure, leaf water content, leaf area index, fractional photosynthetically active radiation absorbed by canopy (FPAR), surface roughness, other foliar chemistry absorption characteristics, and plant phenology) are used as important inputs to a so called radiative transfer model or other biochemical model (Asner and Martin 2008; Li and Strahler 1992; Qin and Liang 2000). By inverting these models, the vegetation parameters of interest can be retrieved numerically. Developments in the physically based modelling approach have shown the model's advantages in robustness and portability over statistical ones (Schlerf and Atzberger 2006; Verhoef and Bach 2007). For instance, radiative transfer models, which describe the interaction of electromagnetic radiation with plant leaves and the canopy using several simple input parameters (Iaquinta et al. 1997; Jacquemoud et al. 2000), have been successfully applied in the forward mode to simulate leaf/canopy reflectance and transmittance of vegetation in general, as well as for inversion to estimate leaf biochemical parameters (Jacquemoud et al. 2009). However, the use of radiative transfer models has been limited to the prediction of plant pigments such as chlorophyll and/or water content. Also, this kind of physical model may lack specificity for tea plants and the special chemical composition in fresh tea leaves.

Statistical or empirical techniques focus on obtaining a statistical correlation between the variables of interest and spectral reflectance values or spectral features of vegetation. Statistical methods help us organize, analyse and interpret data effectively (Kutner et al. 2004; Ott and Longnecker 2010), although the empirical relationships sometimes perform poorly when they are transferred to another location or sensor, or applied over time to the same landscape (Asner et al. 2003; Gobron et al. 1997). In contrast, the physically based models are more robust and portable. However, they too have their disadvantages. Compared to empirical techniques, radiative transfer models have not been used extensively, mainly because they are computationally demanding and require lots of leaf and canopy variables, which are not always easy to estimate (Fang et al. 2003). It is also important to note that the physically based models remain models, that is, simplifications of reality, and require empirical data to weight the model parameters.

Using hyperspectral data from either leaf or canopy, predictive models have been developed for the estimation of some foliar biochemicals based on statistical or physical approaches, or a combination of the two approaches, depending on different aims or conditions (Asner and Martin 2008; Darvishzadeh et al. 2011). To our knowledge, however, little attention has been paid to tea quality assessment using hyperspectral techniques, neither on leaf nor on canopy scale (Dutta et al. 2011).

1.5 Research objectives and questions

The main objective of this thesis is to assess the biochemical parameters of fresh tea leaves associated with tea's quality using hyperspectral remote sensing approaches. I have tried to estimate the concentration of four biochemicals, namely total tea polyphenols, free amino acids, soluble sugars, and caffeine, using spectra from powder, fresh leaves, and the canopy, then explore the possibility of constructing a physically based model for tea quality estimation. The retrievability at different levels, and with different modelling methods, was analysed and compared to explore the potential of hyperspectral remote sensing for assessing the quality of tea on a large scale in the future. The following research questions were formulated to achieve the research objective of this thesis.

⇒ Can the chemical composition (quality-related biochemicals) of tea leaves be detected from spectral reflectance of fresh tea leaves?

⇒ How can the chemical composition of young tea leaves be inferred from spectroscopy of the tea canopy?

⇒ What are the changes in the retrieval capability of spectral bands essential for chemical retrieval from dried tea leaf powder, fresh whole leaves, and canopies?

⇒ Can a physically based model be developed incorporating foliar biochemicals as quality indicators for tea, with the aim to estimate tea quality at a larger scale?

⇒ What is the retrieval accuracy of biochemicals when the physical model is inverted?

1.6 Study site and tea cultivars

1.6.1 Study area

This research was conducted at three different levels – powder, leaf, and canopy level –involving experiments that were carried out in both the laboratory and the field. The laboratory experiment was carried out in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China, while the field survey was conducted in a tea garden located at Huazhong Agriculture University, Wuhan, China (latitude 30°28'41"N, longitude 114°21'48"E). This study area is located in the middle-lower Yangtze plain, and has a subtropical and humid monsoon climate. The average annual precipitation is 1269 millimetres, concentrated between June and August; the average annual temperature is 17.5 degrees Celsius with an annual total of 1950 hours of sunshine.

Part of the data for the study was collected in this tea garden, and part was collected during a greenhouse experiment carried out near the same tea garden (Figure 0-5). A high quality field spectrometer was used to acquire spectra at each level, for both the laboratory and the natural field conditions.

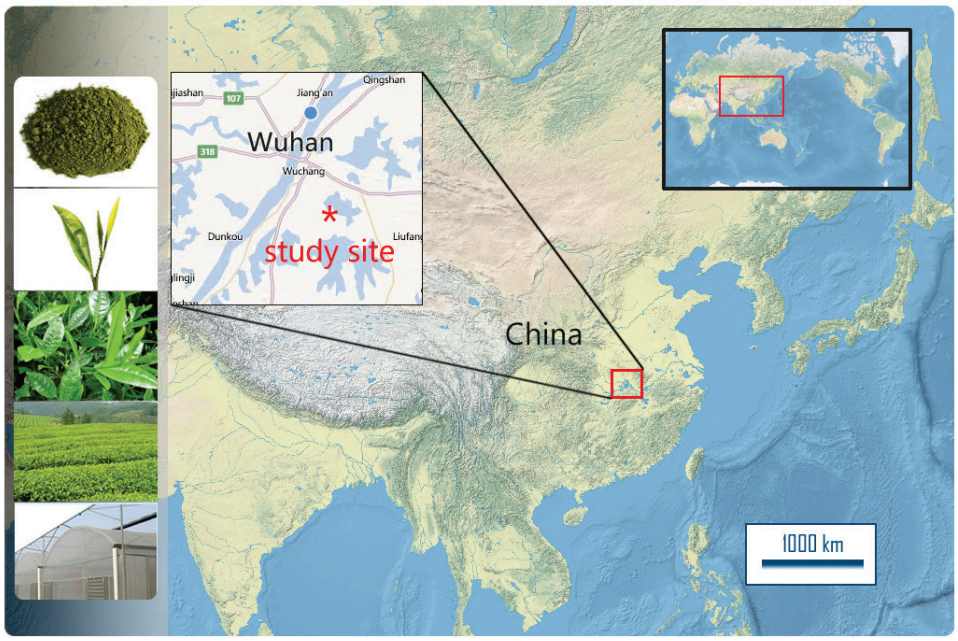

Figure 0-5 Location of the Huazhong Agriculture University, Wuhan, China, where leaf samples were collected from the tea plants cultivated in the tea garden and the greenhouse of the university.

1.6.2 Tea cultivars

There are two major varieties of *Camellia sinensis: sinensis* and *assamica* (Willson and Clifford 1992). They have a different leaf size, shape, and chemical composition. Generally, green tea made in China and Japan is produced from the *sinensis* variety, and the tea leaves are usually picked in early spring, while black tea is made from the *assamica* variety, and the tea leaves picked in summer for high quality. After human cultivation and propagation for thousands of years, many different cultivars of tea now exist (Kumar and Rajapaksha 2005; Liang et al. 1998; Singh et al. 1999). They are usually reproduced through cloning techniques to maintain their flavour characteristics, and used for producing various types of tea. To investigate the spectral-chemical relationship of the tea species, different tea cultivars have been included in the research.

More than ten cultivars of tea are planted in the tea garden of Huazhong Agriculture University for tea cultivation and production experiments. For this

research, mainly the tea variety ‘Fuding dabai’ was used. This tea variety originates from the Fujian province in southern China (Yao et al. 2008). It is now cultivated nationwide for its stable quality green and white teas. Five other Chinese tea cultivars, ‘Fuyun 6#’, ‘E cha 1#’, ‘Tai cha 12#’, ‘Huang dan’, and ‘Mei zhan’, were also used in the study.

1.7 Outline of thesis

The thesis consists of several coherent chapters covering the main objective of the research: to investigate the potential of hyperspectral remote sensing for assessing the quality of tea based on both statistical and physical modelling approaches.

Chapter 1 sketches a general background, states the significance of this research, defines the research objectives, describes the study area, presents the tea cultivars of interest, and outlines the structure of the thesis.

For one tea cultivar, Chapter 2 estimates the biochemical concentrations of total tea polyphenols and free amino acids using both leaf powder and fresh leaf spectra. The hypothesis tested is that the chemical concentration of the two main composition indicators of tea quality can be predicted through hyperspectral techniques at fresh leaf level with reasonable accuracy.

Chapter 3 investigates the spectral variation in tea plants growing in a greenhouse, and links canopy spectra with the chemical composition of fresh tea leaves. An integrated approach, involving artificial neural networks and an effective band selection algorithm, has been proposed to model the possible nonlinear correlation between spectra and biochemicals. The hypothesis tested is that with an up-scale to canopy level, the quality-related biochemicals of tea (total tea polyphenols, free amino acids, soluble sugars, and caffeine) can still be predicted with satisfactory accuracy using hyperspectral data.

Chapter 4 tries to assess the quality of tea using hyperspectral data for mixed tea cultivars planted under natural growing conditions. This chapter systematically investigates and compares the retrieval capability from dried tea powders, fresh tea leaves, and canopies of tea plants. The question of whether the spectral bands essential for chemical retrieval at one scale are still important at other

scales is also answered. Spectral-chemical analysis is carried out for different tea cultivars, trying to detect whether the modelling methods can be extended to various cultivars of the tea species.

Chapters 2, 3, and 4 focus on the estimation of biochemicals as quality indicators of tea using statistical modelling approaches. Chapter 5 proposes an improved leaf optical model for tea (tea-PROSPECT) incorporating total tea polyphenols, aiming to provide more robust and portable results when assessing tea quality across a wide range of conditions in the future. The hemispheric reflectance spectra of fresh tea leaves can be simulated with higher accuracy using the improved physically based tea leaf model, than with the original model which does not use tea polyphenols as separate input parameter. The retrieval accuracy of the total tea polyphenols' concentration by inversion of the tea-PROSPECT model is also evaluated, and compared with the results obtained through empirical regression.

Finally, Chapter 6 links the previous chapters, discusses the interrelationships between these chapters, and proposes further study.

CHAPTER 2

REFLECTANCE SPECTROSCOPY OF BIOCHEMICAL COMPONENTS AS INDICATORS OF TEA QUALITY

This chapter is based on: Meng Bian, Andrew K. Skidmore, Martin Schlerf, Teng Fei, Yanfang Liu & Tiejun Wang, 2010. Reflectance spectroscopy of biochemical components as indicators of tea (*Camellia sinensis*) quality. *Photogrammetric Engineering & Remote Sensing* 76(12), 1385-1392.

Abstract

The potential of reflectance spectroscopy to estimate the concentration of biochemical compounds related to tea (*Camellia sinensis* (L.)) quality (total tea polyphenols and free amino acids) is demonstrated. Partial least squares regression (PLSR) was performed to establish the relationship between reflectance and biochemicals for leaf powders as well as fresh leaves. Highest accuracy was found for tea powders with a cross-validated r^2 of 0.97 for tea polyphenols and 0.99 for free amino acids, and the root mean square error of cross validations (RMSECVs) are 8.36 mg g^{-1} and 1.01 mg g^{-1} for the two chemicals. The accuracy achieved at leaf level was slightly lower, with results yielding cross-validated r^2 of 0.91 and 0.93 with RMSECVs of 13.74 mg g^{-1} and 2.32 mg g^{-1} for tea polyphenols and free amino acids, respectively. Important wavelengths for prediction of the two biochemicals from powder and leaf spectra were identified using the PLSR *b*-coefficients as indicators. Wavelengths of 1131 nm, 1654 nm, 1666 nm, 1738 nm and 1752 nm were identified as bands related to absorption by total tea polyphenols, while 1492 nm represented the absorption feature of free amino acids. The results obtained using fresh leaves indicate that hyperspectral remote sensing may be useful for routine monitoring of tea chemistry at landscape scale.

2.1 Introduction

Tea is one of the most popular beverages consumed worldwide, and is produced from the leaves of young shoots of the plant *Camellia sinensis* (L.). Tea quality is generally judged by its appearance, smell and taste (Chen et al. 2008b; Sharangi 2009; Yan 2007). With the rise in consumption of tea, it has become increasingly important to monitor and control tea quality (Chen et al. 2006).

The quality of tea is determined by the biochemical composition of tea leaves (Obanda et al. 1997; Willson and Clifford 1992) and by the tea processing techniques used (Ravichandran and Parthiban 2000). The processing of tea may be monitored by industrial quality control procedures, while the quality of the raw material is more difficult to monitor and control. Significant relationships have been identified between certain foliar chemical variables and tea sensory preferences (Liang et al. 2008; Obanda et al. 1997). Of the hundreds of chemical compounds found in tea, amino acids and polyphenols are generally considered to be the two main factors determining its quality (Thomas et al. 2009; Yamamoto et al. 1997). The content of amino acids, a major factor in determining the freshness and mellowness of tea, is positively correlated with the quality of green tea. A unique amino acid in tea, named theanine, also produces a relaxation effect in human beings (Chu et al. 1999). Tea polyphenols, 70% of which are catechins, influence the smell and astringent taste of tea and may be of benefit to human health because of antioxidant characteristics (Sharangi 2009). Polyphenols account for 20-35% in weight of the total dry matter of tea.

The quality of tea is traditionally appraised by professional tea tasters, but results may be inconsistent and subjective (Obanda et al. 1997). In contrast, laboratory analysis using wet chemistry techniques is an objective method for detecting biochemical concentrations. Wet chemistry is accurate, but labor intensive. Reflectance spectroscopy offers an alternative method to predict foliar biochemical concentrations, potentially offering fast and low cost analysis at comparable accuracies (Luypaert et al. 2003; Mutanga et al. 2004b).

Reflectance spectroscopy has been used to retrieve the biochemical composition of vegetation from its optical properties since the late 1970s (Curran 1989; Norris et al. 1976) using near infrared spectrometry (NIRS) (Curran et al. 2001).

Molecular vibration of the C-H, N-H, O-H, C-N and C-C bonds within organic compounds cause strong absorptions at wavelengths of 2500-15000 nm (middle-infrared). The overtones and harmonics result in absorption within the near infrared, allowing organic compounds to be detected. Laboratory NIRS has been used to detect alkaloids and phenols, and to determine moisture content in dried tea powder. For instance, Schulz (1999) reported high correlations between tea polyphenols and reflectance spectra. Chen (2008a) predicted the concentration of polyphenols and caffeine with partial least squares (PLS) regression. However, extending the use of NIRS to predict biochemicals at fresh leaf level or, preferably, canopy level has been difficult. The reason for this is that the small absorption features of organic compounds are to a large extent masked by the presence of water and cell structure in fresh leaves (Curran et al. 2001) and biomass (Skidmore et al. 2010).

Imaging spectroscopy could be used to predict tea quality over large areas if stable relations were found between tea spectra and chemicals. Given the advent of hyperspectral sensors, applications of imaging spectroscopy have developed rapidly during the last three decades (Adam et al. 2010; Kokaly et al. 2009; Lee et al. 2007; Majeke et al. 2008; Stagakis et al. 2010; Wessman et al. 1989). Based on hand-held, air-borne and space-borne spectrometers, studies of foliar biochemistry have been carried out for grass (Beeri et al. 2007; Darvishzadeh et al. 2008c), agricultural crops (Peng et al. 2009; Wang et al. 2009), savanna trees (Ferwerda and Skidmore 2007; Skidmore et al. 2010) and forests (Schlerf et al. 2010). In previous studies, chlorophyll, protein, nitrogen, lignin and cellulose in fresh vegetation have been estimated successfully. More recently, foliar macronutrient contents including phosphorous, potassium, calcium and magnesium (Ferwerda and Skidmore 2007), and foliar micronutrients such as sodium (Mutanga et al. 2004a) have been assessed using hyperspectral data. The relationship between hyperspectral data and foliar deterrents such as condensed tannin (Ferwerda et al. 2006a) and polyphenols (Skidmore et al. 2010) has also been modelled. However, linking biochemicals related to tea quality and spectra has not been tested for fresh tea (*Camellia sinensis*) to our knowledge.

This paper aims to estimate the concentration of biochemical compounds that affect the quality of tea, namely free amino acids and total polyphenols, by

using reflectance spectroscopy. Experiments were designed for dried leaf powders and fresh leaves. The objectives are formally to: (*a*) estimate the foliar concentration of total polyphenols and free amino acids using the spectral reflectance of fresh leaves and dried tea powder, and (*b*) identify the most influential wavebands for the estimation of tea polyphenols and amino acids for dried and fresh leaf material.

2.2 Material and methods

2.2.1 Study area

Tea leaves were collected from an experimental field (latitude 30°28'41"N, longitude 114°21'48"E) near Wuhan, which is located in the middle-lower Yangtze plain of China, and has a subtropical and humid monsoon climate. The average annual precipitation is 1269 millimetres, which is concentrated between June and August; the average annual temperature is 17.5 degrees Celsius with an annual total of 1950 hours of sunshine.

The tea named 'Fuding dabai' was chosen for this research as this popular cultivar is widely cultivated. 'Fuding dabai' produces stable quality green and white teas.

2.2.2 Sampling design

Field work was carried out in early June, 2008. Leaf samples were collected at different phenological stages in order to maximize the range of foliar biochemical concentration (Ferwerda et al. 2006a). Using a single season's growth, leaves were picked from a branch, starting at the first leaf with young buds through to the seventh leaf (as young buds are small they were considered to be part of the 'first leaf' sample). For wet chemical analysis leaf samples of at least 70g were required. To generate this weight of leaves, the tea pickers were instructed to continue harvesting leaves from adjacent branches until the threshold of 70g was reached. Each of the seven phenological stages consisted of 8 samples (with each sample being randomly collected from a different part of the tea plantation), resulting in a total sample size of 56. In order to maintain the leaves' freshness, they were stored at 5 degrees Celsius with the time interval from collection to spectral measurement being less than two hours (Cho and Skidmore 2006).

2.2.3 Spectral measurements

Leaf spectral measurements

The collected leaves had their spectra measured using an ASD spectrometer (FieldSpec Pro FR with 25° fibre optics) in a dark-room laboratory. The spectrometer covered a range from 350 to 2500 nm with a spectral resolution of 3 nm for the wavelength interval 350-1000 nm, and 10 nm for the wavelength interval 1000-2500 nm. A re-sampling interval of 1 nm reflectance in 2151 wavelengths was recorded. Each of the 56 samples of fresh leaves comprised a leaf stack that formed an optically dense layer on top of a leaf tray made of thick, black cardboard. The fibre optic cable was pointed at the centre of the leaf stacks at 15º off-nadir, with an instantaneous field of view (IFOV) of 8.9 cm x 8.9 cm. A light source (50W/3200ºK halogen lamp), pointing at the centre of the leaf stacks, was placed at 30º off-nadir. To avoid problems with reflectance hotspots, the sensor and incoming light were placed at a projected angle of 90º (Ferwerda et al. 2006b). Before measuring the spectral reflectance of each sample, a white spectralon panel (Labsphere, Inc.) was used for calibration (Cho and Skidmore 2006). The spectrum of each fresh tea leaf stacks sample was determined as the average of 40 spectral measurements. The leaves were re-mixed after each measurement in order to avoid anisotropy effects and capture the maximum variation in the leaf stack (Cho and Skidmore 2006).

Leaf powder spectral measurements

After measuring the spectra of fresh leaves, the leaves were dehydrated. Each dried leaf stack was divided into two portions. One portion was ground for further biochemical assay (using wet chemistry). The other portion was used for NIRS measurements, which involved grinding and passing the ground leaves through a sieve with a mesh width of 75 μm. The finely ground tea powders were put into a standard container and scanned by a Fourier Transform Near-Infrared Spectrometer (FT-NIR Antaris II) in reflectance mode. The range of spectra was from 1000 to 2500 nm (10,000 cm^{-1}- 4000 cm^{-1}). For each sample unit, the spectral measurements were repeated 3 times. The 3 records were then averaged to form a single spectrum. In the laboratory, both temperature (25 degrees Celsius) and humidity (75% RH) were kept steady.

2.2.4 Biochemical assay

Standard wet chemistry methods were used to determine the concentrations of total polyphenols and free amino acids. The leaves were steamed for three and a half minutes to destroy enzyme activity causing oxidation of the tea before drying in an oven at 80 degrees Celsius. Next, the dried leaves were ground using an electric mill. Total tea polyphenols were determined by the ferrous tartrate colorimetry method and spectrometry at 540 nm (Iwasa and Torii 1962). Concentrations of free amino acids were determined using ninhydrin colorimetry with spectrometry at 570 nm (Curran et al. 2001).

2.2.5 Data analysis

Data Pre-processing

To allow further comparison, the spectra from the NIRS were transformed from log (1/R) to R (reflectance) and interpolated to fit the sampling interval (1 nm) of ASD. Therefore, the NIRS spectra ranged from 1001 nm to 2500 nm and each spectrum contained 1500 wave bands. For the ASD spectra, spectral bands from 350-400 nm and from 2400-2500 nm were excluded from the data because they were considered noisy (based on visual inspection), resulting in 2001 wavebands from 400 to 2400 nm.

Spectral pre-processing may enhance absorption features present in reflectance spectra (Duckworth 1998). There are many factors influencing spectral features such as light scattering, particle size, multi-collinearity among different variables and atmospheric influences. In this study, several spectral pre-processing methods have been applied to reduce these effects, including mean centering, standard normal variate transformation (SNV) and a combination of Savitzky-Golay filtering and first derivative of the reflectance (FDR). Mean centering is a common practice when calculating multivariate calibration models. This involves calculating the average reflectance of all the spectra and subtracting the average from each spectrum (Duckworth 1998). Standard normal variate correction is a method that attempts to remove the major effect of light scattering, and has been used in previous studies to link NIRS spectra to dried tea powders (Luypaert et al. 2003). The derivative spectrum captures the rate of change in reflectance values from band to band, with some studies showing that it improved the prediction accuracy of foliar biochemicals

(Ferwerda et al. 2006b). Before derivative transformation, the spectra were smoothed using a moving Savitzky-Golay filter (Savitzky and Golay 1964). Savitzky-Golay smoothing uses least-square-fit convolution for smoothing and computing derivatives of a consecutive spectrum, and has been regarded as an effective way to eliminate noise from spectra (Ferwerda and Skidmore 2007). The parameters of the fitted second-order, seven-band moving polynomial were applied to reflectance spectra at the center waveband of the moving spline window.

Partial Least Squares Regression

Partial least squares regression (PLSR) has been applied to establish the relationship between concentration of foliar biochemical and reflectance spectra (Cho et al. 2007; Darvishzadeh et al. 2008c). Partial least squares regression combines the features of principal component analysis (PCA) and multiple regressions (Cho et al. 2007). It compresses a large number of variables into a few latent variables (PLS factors). But instead of first decomposing the spectra into a set of eigenvectors and scores, and regressing them against the response variables as a separate step, partial least squares regression model uses the response variable information during the decomposition process. This model effectively reduces the problem of over-fitting found with multiple regression.

In this study, partial least squares regression analyses were performed using the original and transformed spectra, and were processed using the software of ParLes 3.1, which has been developed by Rossel (Rossel 2008). To validate the regression models, a cross validation procedure (leave-one-out cross validation) was adopted. In cross validation, all data in the sample except for a single observation are used for model calibration (training the model), and the model is then tested on the observation that was left out (Schlerf et al. 2005). This procedure iterates until every observation in the training data has been left out. The cross-validated root mean square error (RMSECV) between the predicted and observed values serves as a good indicator of the accuracy in predicting unknown samples (Darvishzadeh et al. 2008a).

The performance of the regression models was evaluated using cross-validated r^2 and cross-validated root mean square error between predictions and observations based on the full data set (n=56). With the purpose of maintaining

model parsimony and avoiding the collinearity problem, the optimal factors for building the model was chosen based on the following criteria: (*a*) the number of factors had to be limited to 1/10 of calibration samples (no larger than 5 in our study) (Marcoulides and Saunders 2006) ; (*b*) an additional factor added to the model had to reduce the cross-validated root mean square error (RMSECV) by larger than 2 percent (Darvishzadeh et al. 2008c).

To determine significant wavebands, PLSR *b*-coefficients are used to indicate the contribution of each predictor (spectral and/or transformed spectral reflectance at each waveband) to the model. In other words, a larger *b*-coefficient (absolute value) indicates a stronger relationship between spectral variables and biochemical concentrations. Gomez et al. (2008) suggested a threshold of one standard deviation (±σ) for *b*-coefficients. Thus, a wavelength was considered significant if the corresponding absolute value of the *b*-coefficients was larger than one standard deviation.

2.3 Results

As detailed in Table 2-1 (n=56), the total polyphenols and free amino acids in the leaves ranged from 146.62 to 294.98 mg g^{-1} and 19.11 to 47.46 mg g^{-1} of the total dry matter, respectively, which matched the normal leaf chemical content of green tea. The coefficient of variation was lower for total polyphenols than for free amino acids.

Table 2-1 Summary statistics for foliar biochemical concentrations of tea; note that all values are reported on a dry-matter basis, and "S.D." represents standard derivation.

Components	*Units*	*Range*	*Mean*	*S.D.*	*Coefficient of Variation*
Total tea polyphenols	mg g^{-1}	146.62 -294.98	192.61	45.86	0.238
Free amino acids	mg g^{-1}	19.11- 47.46	26.54	8.75	0.330

Figure 2-1 (a) shows the mean values of reflectance for fresh leaves from 400 to 2400nm, and for dried powders from 1001 to 2500 nm. The means are flanked by +/- standard deviations. Compared with fresh leaves, powders vary more in reflectance in the SWIR (short wave infrared). In the near infrared and near the water absorption bands, the correlation r of leaf reflectance and powder reflectance decreases from the highest value of 0.85 (around 1650 nm) to

0.6~0.7 (Figure 2-1 (b)), which indicates that for these wavelengths the two types of spectra may contain differing information.

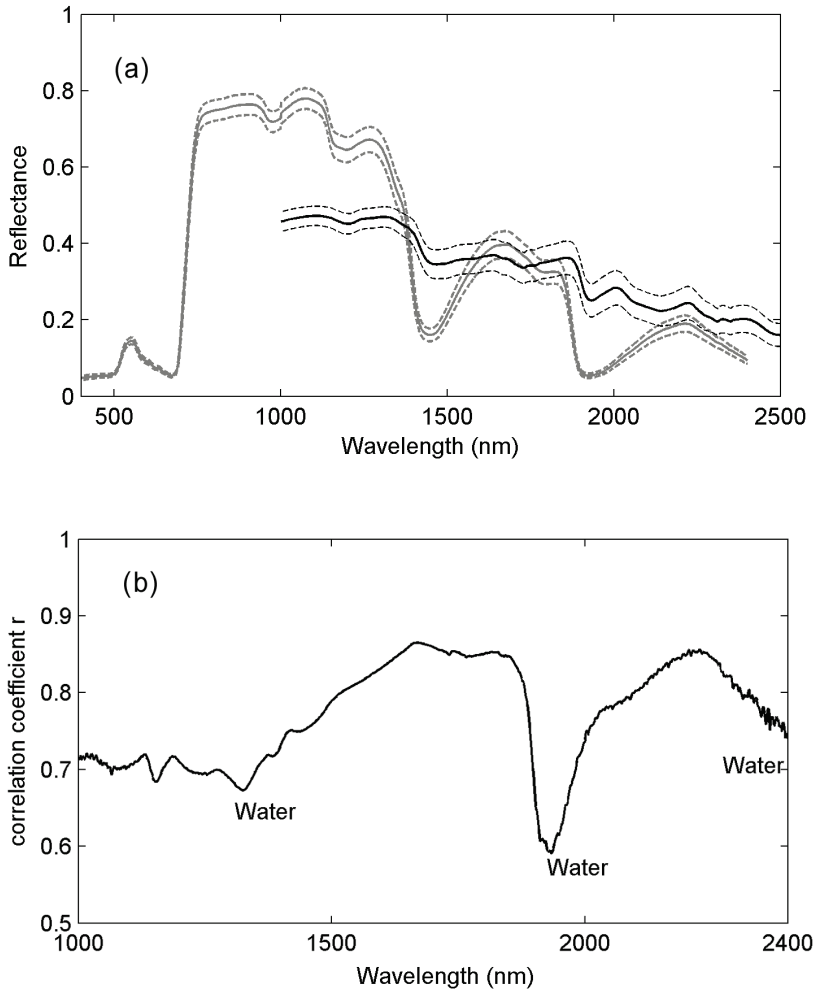

Figure 2-1 (a). Mean reflectance spectra for fresh leaves and dried leaf powders (n=56). The gray solid line represents the mean value of reflectance of fresh leaves; the black line represents the mean value of reflectance of dried leaf powders. The means are flanked by +/- standard deviations (dashed lines), and (b). Correlation between reflected spectra of fresh leaves and dried leaf powders for each spectral band (n=56).

Figure 2-2 shows scatter plots of observed versus predicted concentrations of total polyphenols, calculated using full data set (n=56). Values of r^2 and RMSE are cross-validated. For powders, the highest accuracy of prediction results

(r^2=0.97, RMSECV=8.36 mg g^{-1}) was obtained by a partial least squares regression model using a derivative combined with Savitzky-Golay smoothing. For fresh leaves, the model using spectral prepossessing of SNV provided results with the highest accuracy, although they were less accurate (r^2=0.91, RMSECV=13.74 mg g^{-1}) than those at powder level.

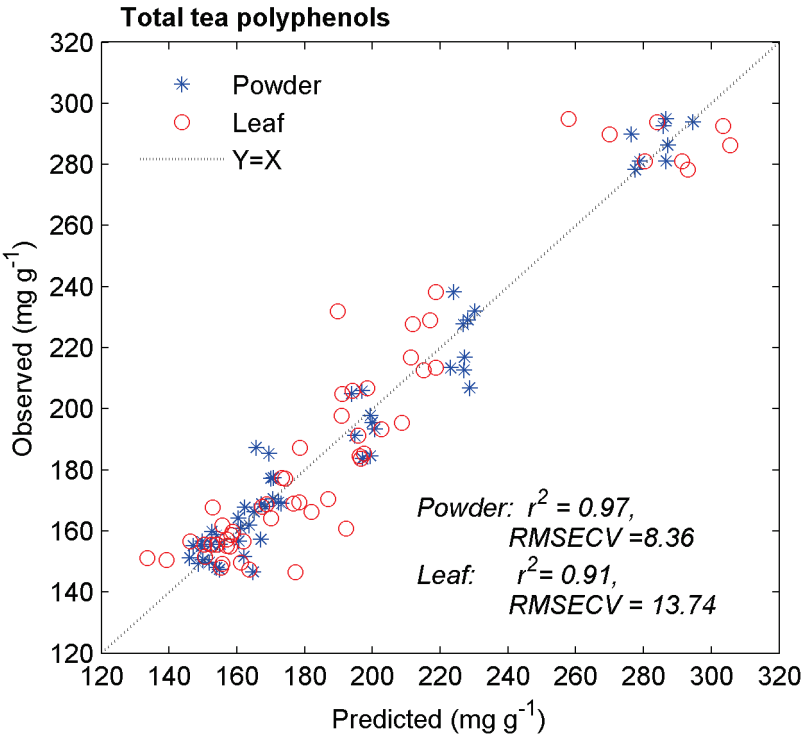

Figure 2-2 Relationship between predicted and observed values of total tea polyphenols for powder (stars) and leaves (circles) using PLSR with the highest accuracy pre-processing method (for powder: first derivative with Savitzky-Golay smoothing; for leaves: SNV). Values of r^2 and RMSE are cross-validated (n=56).

Figure 2-3 details observed versus cross validation predicted concentrations of free amino acids. For powders, the highest accuracy of prediction results (r^2=0.99, RMSECV=1.01 mg g^{-1}) was obtained by a partial least square regression model using first derivative combined with Savitzky-Golay smoothing. For fresh leaves, the PLSR model with inputs of mean cantered spectra gave results with the highest accuracy (r^2=0.93, RMSECV=2.32 mg g^{-1}).

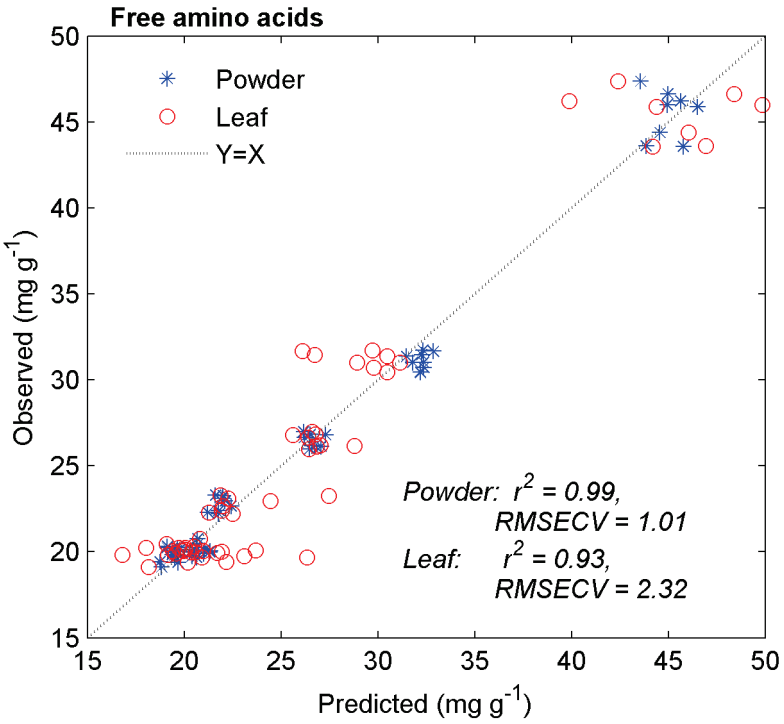

Figure 2-3 Relationship between predicted and observed values of free amino acids for powder (stars) and leaves (circles) using PLSR with the highest accuracy pre-processing method (for powder: first derivative with Savitzky-Golay smoothing; for leaves: mean centering). Values of r^2 and RMSE are cross-validated (n=56).

Table 2-2 details information on the partial least squares regression model when using different pre-processing methods to predict the concentrations of total tea polyphenols and free amino acids for fresh leaves and dried powder. Compared with absolute reflectance spectra, the spectral pre-processing methods increased the prediction accuracy of partial least squares regression models, especially for fresh tea leaves.

Table 2-2 Performance of PLSR with 3 pre-processing methods (pre-processing methods which obtained best results are underlined); note that "RMSECV" represents root mean square error of cross validation and first derivative transformation is performed after Savitzky-Golay smoothing of original spectra.

		Factors No.	Cross-validated r^2	**RMSECV ($mg\ g^{-1}$)	*RMSECV/mean (%)
Total tea polyphenols					
Powders	*No preprocessing*	5	0.96	8.89	4.62
	Mean centering	4	0.97	8.47	4.40
	SNV	5	0.96	8.62	4.48
	<u>first derivative</u>	5	0.97	8.36	**4.34**
Leaves	*No preprocessing*	4	0.87	16.80	8.72
	Mean centering	4	0.90	14.21	7.38
	<u>SNV</u>	4	0.91	13.74	**7.13**
	*first derivative	3	0.89	15.53	8.06
Free amino acids					
Powders	*No preprocessing*	5	0.98	1.29	4.87
	Mean centering	4	0.99	1.04	3.92
	SNV	4	0.97	1.42	5.36
	<u>first derivative</u>	5	0.99	1.01	**3.81**
Leaves	*No preprocessing*	5	0.90	2.77	10.45
	<u>Mean centering</u>	5	0.93	2.32	**8.75**
	SNV	4	0.93	2.34	8.83
	*first derivative	3	0.90	2.71	10.23

Figure 2-4 highlights the influential wavelength channels selected by the PLSR models. It displays the *b*-coefficient change along with the wavelength for the most accurately performing pre-processing method. The parallel lines indicate the thresholds of the *b*-coefficients. Based on the graphs, the continuously important waveband regions and their peaks have been summarized in Table 2-3.

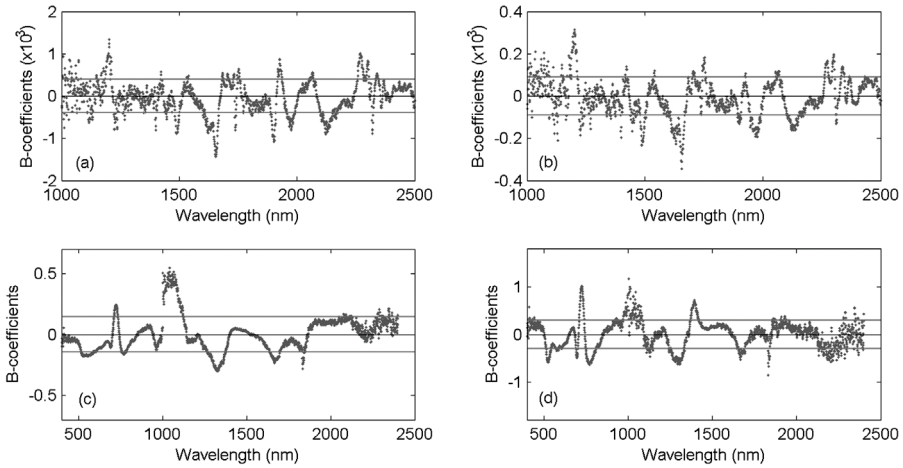

Figure 2-4 *b*-coefficients associated with PLSR models for total tea polyphenols and free amino acids at two levels: (a) for total tea polyphenols in dried leaf powder, (b) for free amino acids in dried leaf powder, (c) for total tea polyphenols in fresh leaves; (d) for free amino acids in fresh leaves. The thresholds for b-coefficients are based on their standard deviation, which are marked using parallel lines.

Table 2-3 Important wavelengths selected by the PLSR model for prediction of total tea polyphenols and free amino acids (start band - peak - end band); note that waveband in bold indicates a causal absorption feature of biochemicals of interest.

	Biochemicals	*Important wavebands (start-peak-end, nm)*
Powders	Total tea polyphenols	1125-**1131**-1133, 1168-1202-1211, 1222-1224-1269, 1482-1488-1499, 1601-**1654**-1667, 1682-1684-1686, 1702-1705-1709, 1736-**1738**-1740, 1747-**1752**-1755, 1888-1902-1909, 1918-1926-1935, 1957-1976-1990, 2060-2066-2069, 2111-2121-2151, 2258-2268-2283, 2294-2301-2305, 2318-2320-2323
	Free amino acids	1148-1202-1213, 1222-1225-1230, 1472-**1492**-1503, 1603-1656-1667, 1743-1753-1759, 1949-1975-1991, 2106-2138-2153, 2262-2274-2281, 2294-2301-2304, 2332-2338-2346
Leaves	Total tea polyphenols	518-553-586, 712-724-734, 759-770-778, 1008-1043-1124, 1267-1326-1375, 1639-**1666**-1695, 1831-1833-1841
	Free amino acids	513-526-541, 691-697-700, 710-720-739, 749-769-809, 1019-1031-1035, 1256-1296-1331, 1371-1393-1418, 1657-1666-1694, 1831-1833-1840, 2178-2182-2186

2.4 Discussion

To the best of our knowledge, this study is the first to accurately predict the concentration of total polyphenols and free amino acids from fresh tea leaves using a high-resolution spectrometer. Our results (Figure 2-2 and Figure 2-3) show that the errors of prediction from cross validation were 13.74 mg g^{-1} (7.13% of the mean) for total tea polyphenols and 2.32 mg g^{-1} (8.75% of the

mean) for free amino acids for fresh leaves. This indicates the potential of reflectance spectra for predicting *in situ* biochemical parameters of tea. It is a step towards the use of airborne and space-borne sensors to monitor tea quality in the field.

The concentration of biochemical components important for tea quality, were predicted with high accuracy from powder spectra (Figure 2-2 and Figure 2-3). This is consistent with previous studies that successfully predicted alkaloids and phenolic substances from dried tea samples using laboratory NIRS. For example, Chen (2008a) measured dried and ground tea samples and found the optimal model for predicting total polyphenols was achieved with a root mean square error of 0.73 (about 4% of mean). Luypaert (2003) explored the possibility to predict the concentrations of EGCG (epigallocatechin gallate, an important component of tea polyphenols) using reflectance spectroscopy. Regarding free amino acids, there has been no previous research using reflectance spectroscopy methods.

Prediction accuracy decreased using fresh leaf spectra, compared to powder spectra. The RMSECV of the total polyphenols from fresh leaf stacks was almost 8% of the mean, compared with 4% for dried powder. For free amino acids, the RMSECV as a percentage of the mean increased from 3.81% to 8.75% between dried powder and fresh leaves. Previous studies indicated that the presence of water in fresh leaves could mask the biochemical absorption features, particularly in the shortwave infrared (Majeke et al. 2008). In addition, by grinding the leaves, the impact of light scattering by intact tea leaves could be minimized or eliminated (Duckworth 1998). Ground powder therefore provided more stable and consistent spectra. Although the errors of prediction increased at leaf level compared to when using tea powders, the retrieval accuracy for both were satisfactory with a RMSECV of less than 10% of the mean.

As shown in
Table 2-3, the most important spectral domains used by the PLSR prediction model were based on the calibration equation coefficients (*b*-coefficients). For NIRS spectra of dried tea powders, spectral features at 1131 nm, 1654 nm, 1738 nm and 1752 nm coincided with total polyphenol concentration as estimated by

PLSR. These wavebands are causally related the earlier reported absorption features of tea phenolic substances. The spectral regions around 1140 nm and 1645 nm have been reported as absorption features for catechins which form the main ingredients of tea polyphenols (Lu et al. 2005), while the spectral features at 1722 nm and 1742 nm have been related to tea polyphenols for dried tea samples (Chen et al. 2006; Chen et al. 2008a). The 1492 nm waveband coincides with the concentration of free amino acids for powders, while Curran (2001) described an absorption feature related to amino acids at 1510 nm. The other wavebands contributing indirectly to total tea polyphenols and free amino acids estimation, are causally located on the absorption features of sugar (1490 nm and 2270 nm), starch (1900 nm and 2320 nm), lignin (1690 nm), protein (1980 nm, 2060nm, 2130 nm and 2300 nm) and water (1200 nm and 1940 nm) (Curran 1989; Osborne and Fearn 1986; Williams and Norris 1987). The concentrations of these biochemicals are correlated with free amino acids and total polyphenols in tea leaves (Yamamoto et al. 1997). Therefore, it is possible that the PLSR algorithm made use of these wavebands to enhance the accuracy of the model prediction.

Fewer influential wavebands were discovered for the spectral features of total tea polyphenols or free amino acids at the fresh leaf level compared to powders. One spectral feature centered near 1666 nm directly contributes to total tea polyphenols estimation, while Soukupova et al. (2002) described an absorption feature related to phenolic compounds at 1648nm. No wavebands directly related to free amino acids contents. This indicates poor signal propagation from the powder to the leaf level. However, PLSR used a lot of indirect bands to build the relationship with chemicals of interest. For example, amino acids and protein are highly correlated and some of the influential bands related to the protein absorption peaks around 1020 nm and 2180 nm were therefore used for amino acids prediction.

The use of imaging spectroscopy for predicting foliar chemistry of tea still faces a great deal of challenges despite the success at leaf level reported in this study. When up-scaling the study from leaf to canopy level, more factors such as canopy structure (e.g. leaf area, leaf angle distribution and canopy closure), soil background properties, illumination and observation geometry and the atmospheric condition can influence the radiance signal received (Kokaly et al.

2009; Majeke et al. 2008). Thus, the detection of biochemical absorption features at the canopy level may be more complicated when using airborne or space-borne hyperspectral data. Also, the optimal wavebands identified from laboratory-based spectra may not be the best from remote sensing platforms.

In order to overcome the challenges of working at the canopy level in future studies, algorithms that minimize the effects of canopy structure, sensor geometry, soil background and atmosphere on canopy spectra must be carefully designed (Mutanga and Skidmore 2004a). An alternative approach is to invert physically based models, for example SAIL (scattering by arbitrarily inclined leaves) model (Verhoef 1984). Using the inverted model, the leaf optical properties can be determined from measured canopy spectra. It is therefore possible to retrieve the foliar chemistry of tea from the simulated leaf spectra, based on the method (partial least squares regression) presented in this paper.

2.5 Conclusions

Overall, the study suggests that total tea polyphenols and free amino acids, as indicators of tea quality, can be successfully estimated using reflectance spectroscopy. The results obtained using fresh leaves indicate that hyperspectral remote sensing may be useful for routine monitoring of tea chemistry at landscape scale. However, to upscale the results from leaf to canopy level and extrapolate the study from laboratory experiment to practice in the field, additional proof and research effort are necessary. Our study also showed that PLSR is an effective method to retrieve biochemical parameters from spectral reflectance of both fresh leaves and leaf powder. PLSR is able to locate surrogate spectral features for estimating the concentration of fresh leaf biochemicals.

CHAPTER 3

ESTIMATION OF TEA QUALITY FROM CANOPY HYPERSPECTRAL MEASUREMENTS

This chapter is based on: Meng Bian, Andrew K. Skidmore, Tiejun Wang, Yanfang Liu & Teng Fei. Estimation of tea quality from canopy hyperspectral measurements. *Food Research International* (in review).

Abstract

Total tea polyphenols, free amino acids, soluble sugars, and caffeine are generally regarded as the major indicators of tea quality. Estimating the concentrations of these biochemical compounds in tea plants before plucking would assist in predicting the quality of tea and allow improvement of tea plantation management. In this study, using hyperspectral data from the tea canopy level, an artificial neural network in combination with a successive projections algorithm was applied to estimate tea foliar biochemical concentrations. The successive projections algorithm was applied to select the optimal wavelength bands from canopy-level hyperspectral data. This integrated approach has led to a satisfactory prediction accuracy: based on an independent test dataset, the best-trained neural networks resulted in a coefficient of determination (r^2) between predicted and observed concentrations of 0.82, 0.76, and 0.76 for tea polyphenols, free amino acids, and soluble sugars, respectively, with a root mean square error of prediction (RMSEP) of 4.30 mg g^{-1} (3.0% of the mean), 1.01 mg g^{-1} (3.7% of the mean) and 3.24 mg g^{-1} (3.5% of the mean), respectively. The accuracy of predicting caffeine was lower with a RMSEP of 1.13 mg g^{-1} (8.4% of the mean). For comparative purposes, other band selection methods including principal component analysis and spectra random selection were also applied to reduce hyperspectral data before running the neural networks, and the new successive projections algorithm proved most accurate. This experiment was carried out in a greenhouse and the biochemical concentration range of the tea was extended by applying different fertilizations. The approach proposed in the study may be extended to a landscape or regional scale, to monitor the quality of tea plantations from airborne or space-borne platforms.

3.1 Introduction

Tea (*Camellia sinensis*) is an important economic crop in China and other developing counties. Besides quantity of tea production, the quality of tea is also important from an economic perspective, as quality is an essential factor influencing the market price (Yan 2007). Factors like colour, appearance, the feel of mouth and smell jointly make up the quality of tea (Dutta et al. 2011). The quality of tea, especially the taste and smell aspects, is largely determined by the biochemical composition of the fresh tea leaves (Obanda et al. 1997; Yamamoto et al. 1997). Thus, it is desirable to develop objective methods to obtain qualitative information in a fast and non-destructive way during the growing period of tea plants.

Determining tea quality is traditionally the task of tea experts, which may lead to inconsistent and subjective results (Acland 1971; Shankar et al. 2003), or, when based on wet chemical analysis, is time and labour consuming. Other physical and chemical methods such as capillary electrophoresis, an electronic tongue, and a lipid membrane taste sensor have been developed to help estimate tea quality. Currently widely available analytical HPLC (high-performance liquid chromatography) method has also been applied for accurate estimation of some biochemicals of tea (Draženka et al. 2010; Wang and Helliwell 2001). However, these methods are still time-consuming and more importantly, they cannot be applied to quality control of tea directly in the field. The development of new techniques using hyperspectral remote sensing data creates the possibility to estimate and monitor vegetation status and quality in space and time in an effective and quantitative way (Clevers et al. 2010; Curran 1989; Knox et al. 2011).

The laboratory-based near infrared spectroscopy (NIRS) technique has been used instead of wet chemical assay for rapid analysis of the chemical content of a wide variety of agricultural and food products (Chen and Lei 2009; Hacisalihoglu et al. 2009; Schulz et al. 1999). The narrow sensitive band range makes it possible to detect subtle variations in the reflectance spectra, caused by differences in biochemical composition of substance (Davey et al. 2009; Ferwerda and Skidmore 2007). In recent years, researchers have extended the technique of reflectance spectroscopy to measuring biochemical parameters of vegetation by field spectrometer, or by airborne or space-borne sensors,

attempting to explain the chemical variation of vegetation in a spatial context (Curran 1989; Schlerf et al. 2010; Skidmore et al. 2010). Bian et al. (2010) estimated the concentration of total polyphenols and free amino acids in fresh tea leaves using a high-resolution spectrometer, based on partial least squares regression. Dutta et al. (2011) measured green and black tea quality in India using a remotely sensed vegetation index and near infrared spectroscopy. But a major drawback in these studies is that the relationship between independent (wavebands) and dependent (biochemicals) variables is assumed to be linear (Farifteh et al. 2007; Mutanga and Skidmore 2004b), though this may not be the case and needs to be explored.

Artificial neural networks (ANN) can model nonlinear relations, and have been regarded as effective methods in the field of chemometrics (Balabin et al. 2008). In vegetation science, the ANN method has been applied to image classification, and recently to the retrieval of biochemical data, such as polyphenols and nitrogen in savannas (Skidmore et al. 2010). For tea quality prediction, Khanchi et al. (2007) used both partial least squares regression and artificial neural networks to determine caffeine and theobromine in Iranian tea fusion based on absorption spectra; Guo et al. (2012) estimated total free amino acids in green tea using near infrared spectroscopy and artificial neural networks with satisfactory accuracy: in the prediction set, the correlation r reached 0.96. However, the application of neural networks to predict tea quality, particularly based on canopy reflectance spectra, has not been studied to date.

Determination of the optimal bands to be assigned to the input neurons of an artificial neural network is one of the critical steps in designing the model (Ferwerda 2005). Data reduction can reduce the processing time without having impact on the performance of the neural networks. Next to other variable selection strategies, such as principal component analysis (PCA), wavelet transform, continuum removal, and genetic algorithms, the successive projections algorithm (SPA) has been proposed as novel and effective variable selection technique (Araújo et al. 2001; Cheng et al. 2009). It is a forward selection method which aims to minimize variable collinearity using simple operations in a vector space. Araújo et al. (2001) applied this algorithm to UV-VIS spectrophotometric data for analysis of chemical complexes, and concluded that the predictive accuracy of the successive projections algorithm was higher

than that of principal component analysis and genetic algorithms. However, for estimating biochemical parameters of plants, SPA has rarely been used as dimension reduction technique for artificial neural network models.

Tea has become increasingly popular in Asian and western countries owing to its possible beneficial effects on human health such as anti-oxidative, anti-tumour and anti-carcinogenic activities which are caused by abundant components such as polyphenols, amino acids and so on (Lin et al. 2012; McKay and Blumberg 2002; Sharangi 2009). Research in the tea industry has also revealed that, among the hundreds of chemical compounds found in tea, total tea polyphenols, free amino acids, soluble sugars, and caffeine are crucial to tea's special flavour and quality. Some tasting tests have identified significant relationships between these foliar chemical variables and human preferences (Kaneko et al. 2006; Liang et al. 2003). Total tea polyphenols is the most biologically active group of the tea components (Venkatesan et al. 2005). The strong astringency of tea is often attributed to its high concentration of tea polyphenols. 70% of tea polyphenols are tea catechins which have a unique astringency that suits the taste of many consumers (Hung, Chen, Chen, & Cheng, 2010). Amino acids form a major factor in determining the freshness and mellowness of tea, and are positively correlated with the quality of green tea. A unique amino acid in tea, named theanine, is primarily responsible for umami taste of tea and has relaxing properties on human being (Chu et al. 1999). The soluble sugars and caffeine are important biochemical components consisting of the water extract which has been regarded as an important international standard for quality control of tea (Yao et al. 2006). Soluble sugars add sweetness and viscosity to tea. Caffeine, the most important alkaloid in tea, contributes to the bitter taste of tea (Dutta et al. 2011; Hui et al. 2002). These above four biochemical components are therefore considered as the important biochemicals influencing tea quality, especially for the green tea (Lin et al. 2012).

The main objective of this study was to develop an integrated artificial neural network and successive projections algorithm for an accurate estimation of quality-related biochemical parameters of tea at canopy level using hyperspectral remote sensing. Total tea polyphenols, free amino acids, soluble sugars and caffeine are the chemicals of interest. A greenhouse experiment was

designed, cultivating 64 groups of tea with 8 different soil treatments, with the aim to discern chemical variations in the plant. This study demonstrates the possibility of an integrated approach to predicting *in situ* tea quality from canopy spectral reflectance. To test the prediction capability of the SPA-ANN algorithm, the results were compared to those obtained by (*a*) the integration of random variable selections and neural networks and (*b*) the integration of principal component analysis (a widely used data reduction method) and neural networks.

3.2 Materials and methods

3.2.1 Plant material and growing conditions

A tea cultivar of Fuding dabai was chosen for this study, because this popular cultivar is widely cultivated in China and produces a stable quality of green and white teas. To save time, it was decided to plant young tea plants (tea seedlings) instead of seeds. In early April 2008, the young tea plants, all cultivated under the same conditions, were transported from a tea plantation in Xianning city in China and transferred in plastic pots (diameter 26 cm, height 20 cm) to a greenhouse, where they were allocated one of eight soil treatments, with eight repetitions. Thus, a total of 64 observations (8*8) were collected for the study. To obtain closed canopies, four tea plants were cultivated in each pot, leaving them sufficient space without competition. In order to have enough leaves for later wet chemical analysis, four or five pots together were regarded as one repetition. A Latin squares design distributed the treatments randomly.

All of the soil treatments were based on the same initial soil composition, to which powdered nitrogen fertilizer ($CO(NH_2)_2$), phosphorous fertilizer (calcium superphosphate) and potassium (K_2SO_4) were added, in either a low or a high concentration (Table 3-1). The high levels of nitrogen, phosphorus and potassium used in this study were adapted from those used in a experiment by Xia (2005), while the low levels were set at one-tenth of the high levels. The fertilizers were applied over a period of three months, in order to avoid excessive fertilization. The different high and low fertilizer combinations resulted in 8 soil nutrient treatments.

Table 3-1 Levels of available soil nutrient used for the 8 treatments (unit: g/pot)

Nutrient level	*Nitrogen*	*Phosphorous*	*Potassium*
Low level (L)	0.75	0.3	0.3
High level (H)	7.5	3	3

3.2.2 Canopy spectral measurements and pre-processing

Measurements were taken after the tea leaf canopy had become dense and there were enough leaves for the wet chemical analyses (after 4 months). The leaves covered the soil, thereby controlling for multiple scattering interaction between soil and vegetation (Mutanga and Skidmore 2004a). An ASD spectrometer (FieldSpec Pro FR) was used for spectral measurements. The spectrometer covered a range of 350-2500 nm with a spectral resolution of 3 nm for the wavelengths 350-1000 nm, and 10 nm for the wavelengths 1000-2500 nm. A re-sampling interval of 1 nm resulted in 2151 wavelengths being included in the measurements. The sensor, with a field of view 10º was mounted on a tripod and positioned about 1.2 m above the ground. On a cloud-free sunny day, the pots were moved out of the greenhouse, and spectral measurements of the tea canopies were carried out between 10:30 and 13:00 hrs.

To avoid bidirectional reflectance distribution function (BRDF) effects (Mutanga et al. 2005), the pots were rotated 60º after every ninth measurement of the canopy. The reflection of a white spectralon panel (Labsphere, Inc.) was measured between readings of the canopy, in order to normalize for differences in irradiance. Because four or five pots were regarded as one repetition, the spectra from these pots were averaged, and the average regarded as one observation.

The band regions 350 nm-400 nm, 1350-1420 nm, 1800-1970 nm and 2300-2500 nm displayed high levels of noise due to atmospheric absorption, and were excluded from the data. The remaining 1657 wavebands were used for analysis. Before data analysis, the reflected spectra of the 64 observations were mean-centered by subtracting their means (Araújo et al. 2001).

3.2.3 Chemical analysis

After taking the reflectance measurements, the tea leaves were collected and dehydrated at 70 degrees Celsius for 24 hours, then stored for analysis. The

concentrations of foliar biochemical compounds were determined using traditional wet chemistry methods in the laboratory. Concentrations of total tea polyphenols were measured using colourimetric method using Folin-Ciocalteu reagent at 765 nm (Ainsworth and Gillespie 2007), according to the standard developed by International Organization for Standardization (ISO 14502-1:2005) and the Chinese new published standard for the determination of biochemical contents in tea (GB/T 8313-2008). Concentrations of caffeine were determined with the UV spectrophotometry method at 274 nm (Ohnsmann et al. 2002). For free amino acids and soluble sugars, concentrations were measured using ninhydrin colorimetry and anthrone colorimetry, respectively (Curran et al. 2001).

3.2.4 Data analysis

Selecting input wavebands based on the successive projections algorithm

The successive projections algorithm (SPA) is a forward selection approach. The purpose of this algorithm is to select wavebands containing least redundant information, so that collinearity problems caused by hyperspectral data are minimized. It starts with one initial waveband *j*. Let X be a matrix of independent variables (spectral reflectance) and X_j be the *j*th column of X. The SPA algorithm then calculates the projection of the reflectance at every other wavelength on the subspace orthogonal to X_j. The wavelength with the largest projection will be preserved as the second informative variable and its projection will be selected as the starting point of the next iteration. The iteration will be stopped when a specified number (N) of wavebands is reached. The projection formula and detailed procedure followed in this algorithm can be found in Araújo et al. (2001).

After spectral pre-processing, the number of input variables was reduced by the successive projections algorithm, with the calculations being performed in MATLAB 7.1 (The MathWorks) software. For each biochemical compound, the optimal number (N) of wavelengths, from a minimum of 1 to a maximum of 31 (the total number of observations in the calibration data set-1), was sought. According to the RMSE values (root mean square error between observed and predicted chemical concentrations based on the data set not employed in the calibration, Eq. 3-1), we found optimal values of N and corresponding wavelength positions for the predictions. With the purpose of avoiding

overfitting and maintaining model parsimony, the criterion for choosing the optimal number of wavelengths was to find the minimum number of wavelengths for which the error was not significantly larger than the lowest error (P>0.05), which was determined by an F-test (Moore and McCabe 1998).

$$RMSE = \sqrt{\frac{1}{AM_{validation}} \sum_{i=1}^{M_{validation}} \sum_{j=1}^{A} [Y'_{validation}(i,j) - Y_{validation}(i,j)]^2} \quad \text{Eq. 3-1}$$

RMSE = root mean square error based on validation dataset
A = the number of analytes (biochemicals)
M = the number of validation mixtures
$Y'_{validation}$ = predicted concentration
$Y_{validation}$ = observed concentration

Selecting the input wavebands based on principal component analysis

Principal component analysis (PCA) is frequently used as a method for data reduction in remote sensing (Bajcsy and Groves 2004; Moore 1981). In this study, the principal component analysis was computed using mean-centred spectra. The first three principal components explained more than 95% of the variability in the spectral data set, and were used as input variables in the neural network to be correlated with chemical concentrations of interest (Schlerf et al. 2005; Schmidt and Skidmore 2003).

Random selection of wavebands (null model)

For each biochemical component, the prediction obtained by using randomly selected wavebands as input data in the neural networks was also computed to compare the quality of the successive projections algorithm solution with the solution obtained by randomly selecting wavelength bands. The calibration ANN model with a random selection solution was defined as null model. Taking calculation time into consideration, we randomly selected 1-20 wavebands over 1000 iterations. The error distribution of the 1000 iterations was compared with the results from the successive projections algorithm.

Artificial neural networks

ANNs are statistical optimising and learning methods. When the feature space is complex, neural networks may provide more accurate results than other statistical methods (Zhang and Friedrich 2003). The back propagation method is

a learning procedure for multilayered, feed-forward neural networks. It was one of the most widely used training methods for ANNs. The structure of a back-propagation neural network (BPNN) consists of several layers; each layer is connected with weights and bias. Learning is carried out by iteratively adjusting the weights and bias in the network so as to minimize the differences between the actual output of the network and the target output. We used a back-propagation neural network of an input layer, a hidden layer, and an output layer, to assess the biochemical concentrations of tea (Figure 3-1). The wavebands selected by the successive projections algorithm, by principal component analysis, and by random selection were each used separately as input in the neural network. The prediction accuracy achieved by the successive projections algorithm solution was compared with the results obtained using the neural network in combination with the other two band selection methods.

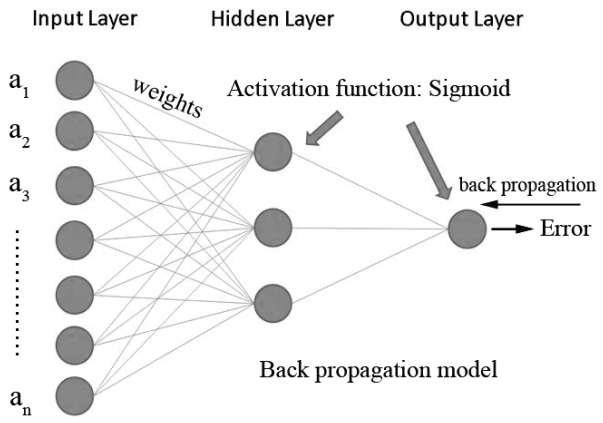

Figure 3-1 The conceptual architecture of a three-layer back propagation artificial neural network.

Being a stochastic method, the performance of a neural network depends strongly on the optimal operational settings, such as number of neurons in the hidden layer and the combination of learning rate, momentum, and number of iterations (epochs) (Mutanga and Skidmore 2004b; Skidmore et al. 2010). To find the optimal number of nodes in the hidden layer, we investigated the training and test accuracies using different numbers of neurons (1-20) in the network (the maximum number was limited to 20 to maintain model parsimony

and save calculation time). An early stopping technique (Lin and Chen 2004) was applied in this study to avoid overtraining. This technique divides the available data into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation dataset. Besides the training error, the error on the validation set is also monitored during the training process. The validation error and the training error normally both decrease during the initial phase of training. However, when the network begins to overfit the training data, the error in the validation dataset typically begins to rise. When the increase of error is detected, the training is stopped, and the weights and biases at the minimum of the validation error are returned (Coleman et al. 1999). In this study, the proportions of the three data subsets are: training data (n=32), 50% of the available data; validation dataset (n=16), 25% of the available data and validation dataset (n=16), 25%. A Levenberg-Marquardt optimisation method was used to train the networks(Lera and Pinzolas 2002). The Levenberg-Marquardt optimisation method is a blend of local search properties of Gauss–Newton with consistent error decrease provided by gradient descent algorithm. The training of feed forward network based on Levenberg-Marquardt method is considered to be an unconstrained optimisation problem. This method was designed to approach a second-order training speed without having to compute the Hessian matrix. Since the learning rate in this method is not a constant value but to be updated using a "decay rate" which depends on the outcome error of each iteration, the learning rate was not specified. Instead, following the parameter suggested in (Suratgar et al., 2005), an arbitrary initial value of 0.01 for the learning rate was assumed appropriate.

For each biochemical component, different ANN architectures with 1 to 20 neurons in the hidden layer were attempted. For each architecture, the training process was run ten times with randomized initial weights. The prediction capability of the neural network was assessed using the coefficient of determination r^2 and RMSEP values between observed (wet chemical analysis) and predicted (i.e. derived from the neural network) chemical concentrations, and calculated using the independent test dataset (Skidmore et al. 2010; Skidmore et al. 1997). By comparing the averaged prediction accuracies of every ten runs, the optimal numbers of hidden nodes producing the highest r^2 as well as the lowest RMSEP were chosen for the prediction of tea polyphenols,

free amino acids, soluble sugars, and caffeine.

3.3 Results

3.3.1 Descriptive statistics of foliar chemistry

Table 3-2 shows the average concentration of tea polyphenols, free amino acids, soluble sugars, and caffeine of tea plants after four months growth and for 8 different soil nutrient levels. All values are reported on a dry-matter basis. In general, the highest level of fertilizer application ($N_HP_HK_H$, see Table 3-2) resulted in the highest concentration of plant foliar biochemicals, though, for example, soluble sugar had the highest concentration with the fertilizer combination $N_HP_LK_H$, and caffeine with the fertilizer combination $N_HP_HK_L$. The range of the chemical data is in accordance with previously reported values (Yamamoto et al. 1997).

Table 3-2 Average foliar biochemical concentration of tea plants for different soil nutrient levels (the minimum and maximum values are shown respectively in brackets). N_H, P_H, and K_H denote high levels of nitrogen, phosphorus and potassium, respectively; N_L, P_L, and K_L denote low levels of nitrogen, phosphorus and potassium, respectively.

Soil treatment	*Repetitions No.*	*Tea polyphenols (mg g^{-1})*	*Free amino acid (mg g^{-1})*	*Soluble sugars (mg g^{-1})*	*Caffeine (mg g^{-1})*
$N_LP_LK_L$	8	**126.28** (118.40-134.62)	**25.75** (25.16-26.73)	86.06 (82.17-87.38)	13.47 (13.28-13.88)
$N_LP_LK_H$	8	132.36 (126.57-138.42)	25.76 (23.56-27.58)	**84.34** (82.42-86.18)	**10.46** (10.31-10.56)
$N_LP_HK_L$	8	132.87 (125.77-137.27)	27.36 (25.83-28.54)	84.68 (82.45-87.17)	13.97 (13.77-14.13)
$N_LP_HK_H$	8	133.83 (129.83-141.35)	27.73 (26.75-29.22)	88.18 (87.23-90.51)	12.10 (11.88-12.41)
$N_HP_LK_L$	8	132.59 (127.32-138.98)	29.07 (25.39-32.74)	93.15 (91.10-94.60)	13.98 (13.95-13.41)
$N_HP_LK_H$	8	143.88 (137.02-149.53)	28.11 (27.54-28.48)	**98.30** (95.77-101.25)	13.62 (12.61-14.46)
$N_HP_HK_L$	8	133.24 (125.73-141.01)	29.41 (28.44-30.22)	97.68 (96.24-99.15)	**16.20** (14.70-17.27)
$N_HP_HK_H$	8	**145.96** (141.89-149.99)	**29.63** (27.28-34.64)	92.99 (90.79-95.21)	16.04 (14.85-17.71)

3.3.2 Training and testing the neural networks based on the wavebands selected by the successive projections algorithm (SPA-ANN)

Figure 3-2 displays the optimal choice of wavelength selected by the successive projections algorithm. The best choice of 12, 13, 7 and 17 wavebands has been selected for the prediction of total tea polyphenols, free amino acids, soluble sugars and caffeine, respectively. The selected wavebands were used as input for the neural networks.

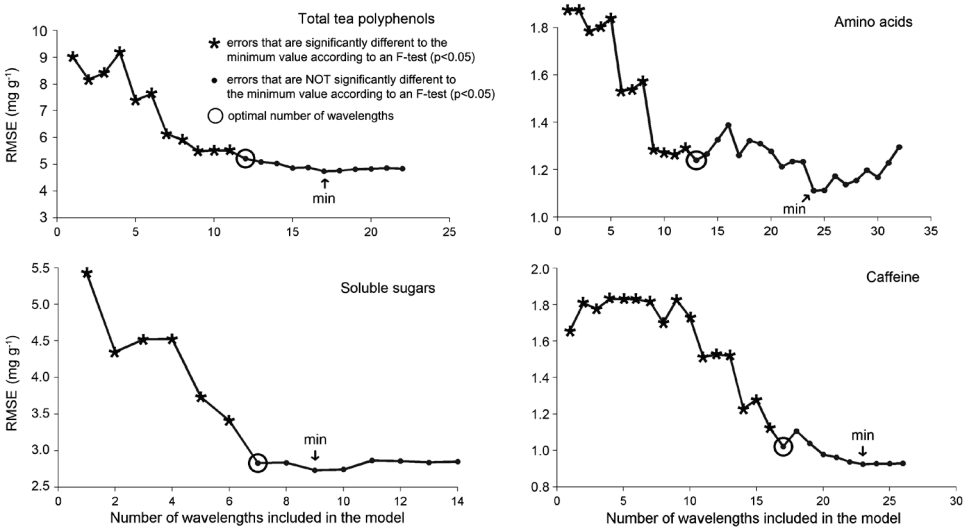

Figure 3-2 Choice of the optimal number of wavelengths (circled positions) by successive projections algorithm for the prediction of the biochemicals of interest. The criterion is to find a minimum number of wavelengths for which the error is not significantly larger than the lowest one (P>0.05), as determined by an F-test.

Figure 3-3 presents the relationship between predicted and measured biochemical concentrations using a hybrid classifier comprising a neural network and input from the successive projections algorithm variable selection (SPA-ANN). Using the wavebands selected by the successive projections algorithm, the neural network (with optimal settings derived as described above) yielded a coefficient of determination r^2 of 0.82, 0.76, 0.76, and 0.74, respectively, for the prediction of total tea polyphenols, free amino acids, soluble sugars, and caffeine for the test data set (n=16), with a root mean square error of 4.30 mg g^{-1} (3.0% of the mean), 1.01 mg g^{-1} (3.7% of the mean), 3.24

mg g^{-1} (3.5% of the mean), and 1.13 mg g^{-1} (8.4% of the mean), respectively. In other words, prediction of caffeine using the SPA-ANN model was the least accurate. Figure 3-4 shows training and test accuracies in response to the changing number of hidden nodes (1-20) in the neural network for each biochemical prediction. According to the minimum RMSEP values based on the independent test data, the chosen optimal neural network model for the total tea polyphenol concentration had 7 neurons in the hidden layer of the ANN, for the free amino acids this was 18 neurons, for the soluble sugars 6 neurons, and for caffeine 15 neurons.

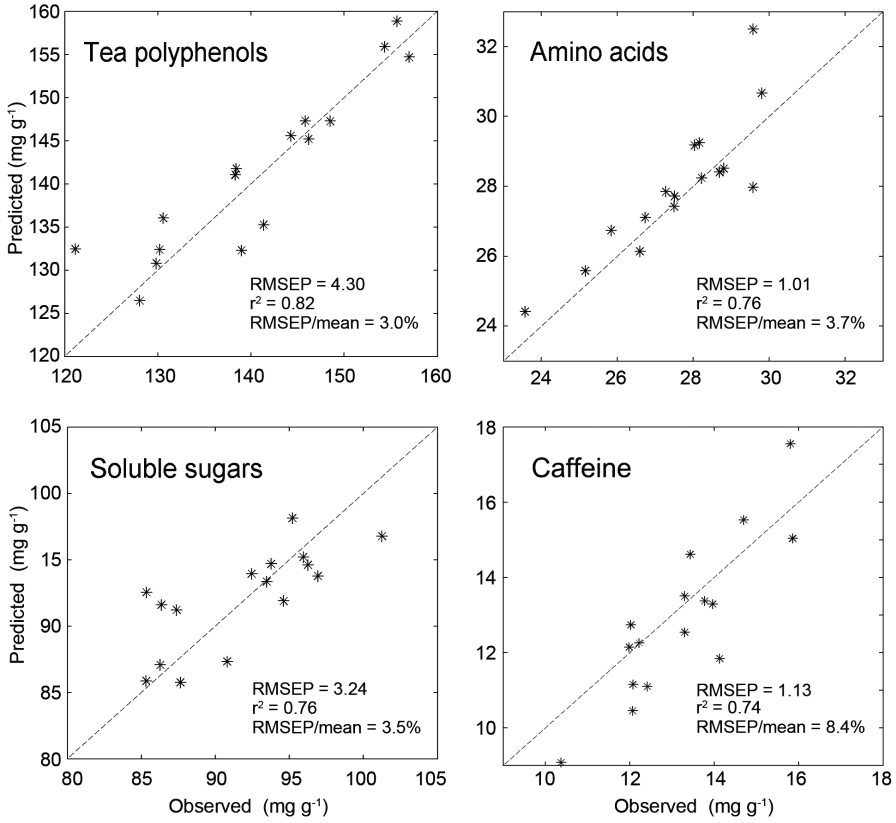

Figure 3-3 Relationships between the predicted and measured biochemical concentrations using a hybrid of neural networks and SPA variable selections, based on the test dataset (n=16). The result was derived from optimal neural networks with the number of hidden nodes selected on the basis of the highest test prediction accuracy.

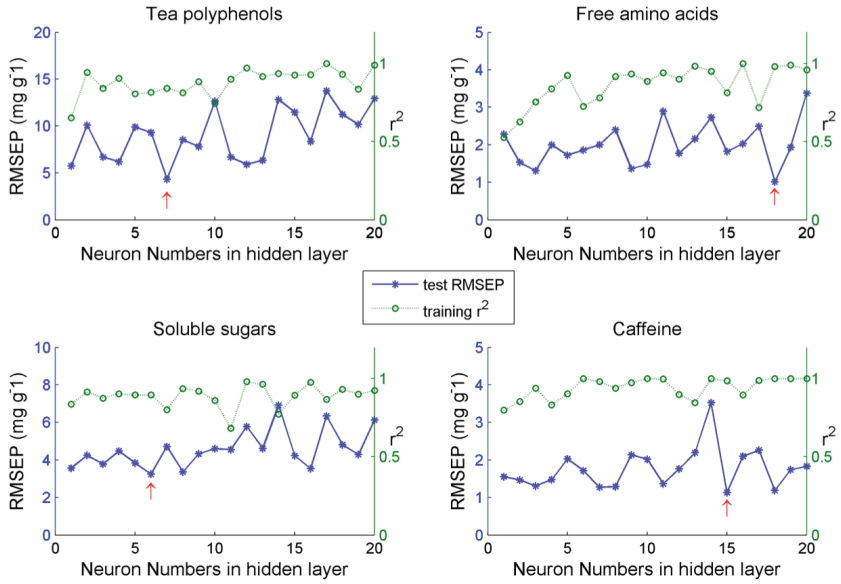

Figure 3-4 Training (circles) and test (solid dots) accuracies in response to the changing number of hidden nodes (1-20) of neural networks. For each biochemical, the minimum RMSEP value is indicated by an arrow and the corresponding number of neurons has been selected as optimal (viz. 7 for tea polyphenols, 18 for amino acids, 6 for soluble sugars, and 15 for caffeine). Each plot in the figure represents the average prediction accuracy of 10 iterations with certain number of neurons.

3.3.3 Comparison of models with SPA, PCA and random selections

Comparing the above successive projections algorithm results with principal component analysis (the first three principal components), the highest accuracy was achieved with the SPA-ANN models (for the prediction of all the biochemicals of interest), as detailed in Table 3-3. When compared with the randomly selected wavelengths, the SPA solution outperformed most of the random band combinations and yielded higher prediction accuracy. Figure 3-5 demonstrates the prediction accuracy of the neural network in combination with all three waveband selection strategies, i.e. successive projections algorithm, principal component analysis, and random selection of wavebands, based on independent test data. The random selections were iterated 1000 times and generated a distribution of RMSEP, which is presented using box plots.

Table 3-3 Comparison of the variable selection solutions of the successive projections algorithm and principal component analysis.

			Independent test data set (n=16)		
	Number of inputs	Number of neurons	r^2	RMSEP* (mg g^{-1})	RMSEP/mean (%)
Total tea polyphenols					
*SPA-ANN***	12	7	0.83	4.30	**3.03**
*PCA-ANN****	3	6	0.50	5.90	4.28
Free amino acids					
SPA-ANN	13	18	0.76	1.01	**3.72**
PCA-ANN	3	8	0.50	1.14	4.00
Soluble sugars					
SPA-ANN	7	6	0.74	3.24	**3.47**
PCA-ANN	3	5	0.48	4.21	4.68
Caffeine					
SPA-ANN	17	15	0.76	1.13	**8.39**
PCA-ANN	3	4	0.32	1.40	10.3

* Root mean square error of prediction using test data set, ** Neural network in combination with successive projections algorithm, *** Neural network in combination with principal component analysis.

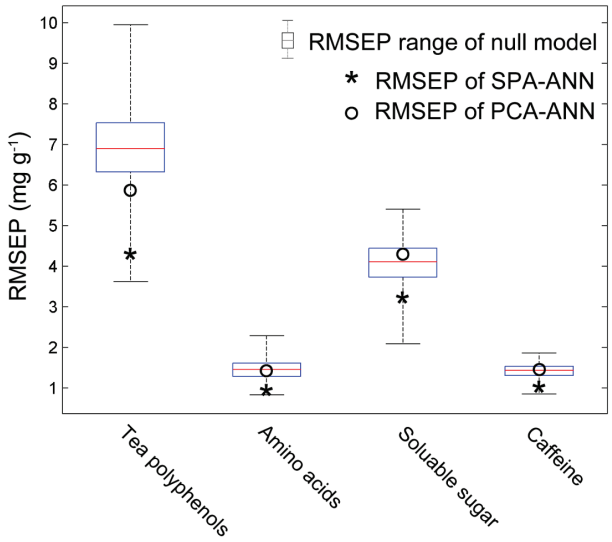

Figure 3-5 Prediction accuracy yielded by the ANN model in combination with the successive projections algorithm (RMSEP of SPA-ANN, asterisks), principal component analysis (RMSEP of PCA-ANN, circles), and random selections of wavebands (RMSEP of null model, boxplots) (n=16). The boxplots shows the upper/lower limits, the 50% quartiles and the mean values for the 1000 random iterations.

3.4 Discussion and conclusions

This study has demonstrated that the quality of tea can be predicted with satisfactory accuracy using hyperspectral data processed by an artificial neural network in combination with the successive projections algorithm. Based on the optimal number of wavelengths selected by the successive projections algorithm, the neural network accurately predicted the concentration of tea polyphenols, free amino acids, soluble sugars, and caffeine: the relative root mean square errors (RMSEP/mean) were less than 10% on an independent test dataset (Figure 3-3). In previous studies, laboratory near-infrared spectrometry (NIRS) has been used to predict quality related biochemicals of tea from dried tea powders (Luypaert et al. 2003; Schulz et al. 1999). Recently, Bian et al. (2010) extended the use of reflectance spectroscopy to predict biochemicals for fresh tea leaves. Here results were achieved at tea canopy level, highlighting the potential for hyperspectral imagery to predict the quality of tea over a large area.

To predict the four chemical components of interest, the lowest root mean square error of prediction was obtained using the successive projections algorithm to select the band variables (Figure 3-5 and Table 3-3). The successive projections algorithm appears to optimise the number of inputs and maintain the parsimony of the neural network. In other words, the goal of this algorithm is to find a small representative set of spectral variables while minimizing collinearity. These results are consistent with those of recent studies which successfully predicted moisture and biochemical concentrations in wheat and rape leaves using the successive projections algorithm as spectral variable selections (Cheng et al. 2009; Liu et al. 2010).

The predictive performance of the neural network in combination with principal components analysis was higher than with random selections, but could not compete with the accuracy of the successive projections algorithm. For the decomposing process of principal component analysis was a stand-alone step without considering the response variable information, the result may lack of specificity and lead to a lower accuracy (Cho et al. 2007).

Comparing the results obtained by successive projections algorithm to those obtained by random selection, we found that the band selection largely

improved the predictive performance of the neural network for total tea polyphenols, free amino acids, soluble sugars, and caffeine (Figure 3-5). In our study, predicting caffeine using a neural network obtained a lower accuracy: with the successive projections algorithm, the relative root mean square error was 8.4% for the test data. This weak predictive power may be caused by the relatively low concentration of caffeine in tea plants, namely 10.31 mg g^{-1} to 17.71 mg g^{-1} (less than 2% of the dry matter) (Cho and Skidmore 2006).

In our study, young tea plants were grown in a greenhouse experiment on soils with different levels of available nitrogen, phosphorus and potassium. The high levels of nutrients applied were adapted from those used in a pot experiment by Xia et al.(Xia et al. 2005), also it was consistent with the upper limit of fertilization for each tea plant in field in tea plantations. Moreover, in our study the canopy reflected spectra were measured from four young tea plants together for each pot. This better resembles the *in situ* situation of tea plantations where usually several (about three) tea seedlings are planted as one tea cluster. Although we have young tea plants in this greenhouse experiment, in the field there are generally aged ones. As some sources in the literature have pointed out, multiple scattering in the upper canopy leaf layers might enhance the detect ability of leaf biophysical properties (Schlerf et al. 2010). In particular this holds true for highly foliated canopies of aged tea plants where leaf area index is larger than 5 (Asner and Martin 2008). Thus, the approach proposed here can be extended to further research in the field where more reliable in situ tea quality information can be retrieved by imaging spectroscopy. However, we should also note that as this experiment was undertaken in a greenhouse under controlled conditions, only the different soil nutrient concentrations affected the biochemical concentrations (Table 3-2), causing the subsequent spectral variation. When using airborne or space-borne hyperspectral remote sensing, the retrieval of biochemical parameters for tea plants may be more difficult, as the biochemical absorption features will be affected by complex environmental factors (Kokaly et al. 2009) or interact with other plant variables such as biomass (Dabrowska-Zielinska et al. 2009; Skidmore et al. 2010).

In summary, this paper proposes a novel integrated approach, involving a forward selection algorithm (successive projections algorithm) to choose the optimal number of wavelengths and neural network settings, to predict

biochemical concentration as a quality indicator for tea. As a variable selection strategy, the successive projections algorithm successfully reduces the large number of wavelengths in hyperspectral data, and outperforms the widely used dimensionality reduction method of principal components analysis. The successful chemical estimation from canopy spectra demonstrates the possibility of using hyperspectral remote sensing to predict tea quality quantitatively and non-destructively in space and time before its plucking.

CHAPTER 4

PREDICTING FOLIAR BIOCHEMISTRY OF TEA USING HYPERSPECTRAL DATA FROM POWDER, LEAF TO CANOPY

Meng Bian, Andrew K. Skidmore, Martin Schlerf, Tiejun Wang, Yanfang Liu, Rong Zeng & Teng Fei. Predicting foliar biochemistry of tea using hyperspectral data from powder, leaf and canopy. *ISPRS Journal of Photogrammetry and Remote Sensing* (accepted).

Abstract

The concentrations of biochemical compounds (i.e., total tea polyphenols, free amino acids and soluble sugars), those are related to the quality of tea across different tea cultivars, were estimated using partial least squares regression, as based on measurements by reflectance spectroscopy at three different levels (i.e., ground powder, fresh leaf and canopy). This is the first study that systematically compares the predictive power of tea chemistry models for ground and dried leaves, whole fresh leaves and tea plants. Using independent test data set, at the powder level, the average r^2 between predictions and observations were 0.89 for polyphenols, 0.81 for amino acids and 0.78 for sugars, with the relative root mean square errors (RMSE/mean) of 5.47%, 5.50% and 2.75%, respectively; at the leaf level, the average r^2 decreased to 0.46-0.81 and the relative RMSE increased to 4.46%-7.09%. Compared to the results yielded at the leaf level, the results from canopy spectra was slightly better, which yielded average r^2 of 0.83, 0.77 and 0.56 and relative RMSE of 6.79%, 5.73% and 4.03% for polyphenols, amino acids and sugars. We further identified wavelength channels that influenced the prediction model. For powders and leaves, some bands identified as important can be linked to the absorption features of chemicals of interest (1648 nm for phenolic, 1510 nm for amino acids, 2080 nm and 2270 nm for sugars), while more indirectly related wavelengths were found important at the canopy level for the prediction of chemical compounds. Overall, the prediction accuracies achieved at canopy level in this study are encouraging for future tea quality estimates at the landscape scale using airborne and space-borne sensors.

4.1 Introduction

The use of reflectance spectroscopy to detect some foliar biochemicals has been developed, tested and proven effective for dried, ground samples such as protein, lignin and starch concentrations under controlled laboratory conditions (Curran, 1989; Elvidge, 1990; Ferwerda and Skidmore, 2007). As early as the 1970s, using regression calibration methods, the statistical relationships between 42 absorption features in the visible and near-infrared wavebands and the concentrations of organic compounds (e.g., cellulose, lignin, protein, oil, sugar, starch) in dried leaves had been established by researchers from the United States Department of Agriculture (Curran, 1989; Norris et al., 1976). However, most of the relationships are statistical correlations, with the exception of a few studies that have tried to explore the physical relation linking foliar chemical contents with their spectra (Curran, 1989; Curran et al., 2001; Kokaly et al., 2009). To obtain more accurate results, spectral transformations such as first derivative (Wessman et al., 1988a), band-depth analysis (Curran et al., 2001; Kokaly and Clark, 1999) and principal component analysis were applied as pre-processing steps. The results showed that for a wide range of biochemicals, spectral pre-processing methods enhanced the accuracy of the estimations (Duckworth, 1998; Hauksson et al., 2001; Joffre et al., 1992; Schulz et al., 1999). Today, reflectance spectroscopy has become a standard method that enables quick, accurate estimations of certain plant biochemicals such as chlorophyll and nitrogen in many laboratories (Wessman et al., 1988a).

Besides successful applications in estimating biochemicals from dried foliar samples using reflectance spectroscopy, remote sensing of biochemical detection has shown great potential (Card et al., 1988; Kokaly et al., 2009; Martin and Aber, 1997; Wessman et al., 1988b). The advent of hyperspectral remote sensing (in the 1980s) has made it possible to detect plant biochemicals at a large scale, for the large number of bands and the narrow sensitive range enable the detection of changes in narrow absorption features which is not possible with broadband data (Blackburn, 1998; Galvão et al., 2008; Song et al., 2011). For example, Skidmore et al. (2010) mapped the concentrations of foliar protein and polyphenols in trees and grass in savanna using airborne hyperspectral imagery. In recent years, hyperspectral remote sensing technology has increasingly been used for the estimation of leaf biochemistry in forage (Darvishzadeh et al., 2008c; Mutanga and Kumar, 2007), cultivated crops

(Haboudane et al., 2002; Hansen and Schjoerring, 2003) , forests (Asner and Martin, 2008; Schlerf et al., 2010) and other tree species .

From dried leaf powder, fresh whole leaves to canopies, the complexity in spectroscopy typically increases. It is easier to observe absorption features from powdered leaves than from fresh leaves, because the water in fresh leaves has strong absorption features and may mask the minor absorption features of organic compounds (Curran et al. 2001; Peterson et al. 1988; Ramoelo et al. 2011b). Compared to fresh leaf, the canopy is even more complicated as factors such as soil background and canopy structure are introduced to the measured signal. Although previous studies have made substantial progress in retrieving foliar biochemistry at each scale, few studies have been done to compare the retrieval capability at the three different scales. For instance, for tropical forest, Asner and Martin (2008) detected chlorophyll, carotenoids, foliar nutrients and water using leaf and canopy spectra, and compared the spectral-chemical correlations between the two scales in a tropical forest. However, it remains unclear whether the spectral bands essential for chemical retrieval at one scale are still important at other scales.

Stepwise multiple regression (SMR) has been used to predict foliar nutrients from spectra. However, it is likely to suffer from multi-linearity (Cho et al. 2007; De Jong et al. 2003). Partial least squares regression (PLSR), an improved multivariate statistic technique can deal with this problem (Duckworth 1998). It compresses a large number of variables (wavebands) to a few latent variables (PLS factors) (Cho et al. 2007). At the same time, based on regression coefficients, bands who contribute most to the prediction can be identified (Gomez et al. 2008). Because of these advantages, partial least squares regression method has been applied to predict vegetation parameters (Cho et al. 2007; Darvishzadeh et al. 2008c; Huang et al. 2004) and was applied in this study.

Tea plants have been cultivated for millennia. The selection process of varieties has been limited to those plants (clones) which can produce an acceptable cup of tea, thus shaping the chemical characteristics of today's tea (Agarwal et al., 1992; Banerjee, 1988). Fresh tea leaves contain caffeine, tea polyphenols, tea polysaccharides, and necessary nutrients, such as protein, amino acids, lipids,

and vitamins. Generally, some chemical components – free amino acids, total tea polyphenols, and soluble sugars – are considered important indicators of tea quality (Ruan et al., 2010; Yamamoto et al., 1997; Yuan, 2003). The special flavor and astringency of a tea brew is the result of the total tea polyphenols (including catechin, epigallocatechin-3-gallate (EGCG), epigallocatechin epicatechin-3-gallate, epicatechin and their epimers), which amount to 18-36% of the weight of the total dry matter of a tea leaf (Graham, 1992). They easily oxidize and may react with some other chemical constituents during the processing, thereby influencing the taste of tea infusion (Obanda et al., 2001). The freshness and brothy taste of tea originates in the free amino acid fraction that is soluble in water, especially the unique amino acid in tea, theanine (L-theanine), which accounts for more than 50% of the total amino acids in tea leaves (Willson and Clifford, 1992). Soluble sugars which amount to about 10% of the weight of the dry matter of a tea leaf add sweetness and viscosity to tea infusion (Ding et al., 2002; Scharbert and Hofmann, 2005).

This research aims to estimate the chemical compounds related to tea (Camellia sinensis (L.)) quality, including total tea polyphenols, free amino acids and soluble sugars, using reflectance spectroscopy at three different scales - ground powder, fresh leaf and canopy. Partial least squares regression was performed to establish relationships between reflectance and biochemical contents, because it has proven to be the one of the most successful empirical approaches for deriving foliar chemicals from canopy spectra (Kokaly et al., 2009; Ramoelo et al., 2011). The specific objectives are: 1) to establish empirical relationships between the concentrations of key chemicals and the reflectance spectra at three different levels; 2) to assess the accuracy of the established models in predicting unknown sample concentrations and to evaluate the prediction accuracy among the different levels; 3) to determine the important wavebands at each level, and further assess if the important wavebands derived from ground powder spectra coincide with the wavebands derived from fresh leaf/canopy spectra.

4.2 Methods and materials

4.2.1 Study area

The study was conducted at an experimental tea field in Wuhan city of China (30°28'41"N, 114°21'48"E). Wuhan was located in the middle-lower Yangtze plain. Average annual temperature is about 17.5 degrees Celsius. The annual

sunshine time is greater than 1800 hours. The average annual rainfall is about 1,200 mm and mainly concentrated in June, July and August. More than ten different cultivars of tea are planted in this tea garden for breeding and production experiments. The tea bushes are dense so dense, in fact, less than 1% of the soil beneath them in the field of view can be seen due to canopy cover.

4.2.2 Sampling design

A sample of 48 observations was designed for the study. Spectral measurements included the spectral reflectance of the tea at powder, fresh leaf and canopy levels; a wet chemistry assay included the concentration of chemical compounds of tea polyphenols, free amino acids and soluble sugars in the tea leaves. To ensure enough leaves could be collected for the wet chemistry assay (which required at least 10g of dried leaf powder per sample point, equivalent to 50g of fresh leaves), a group of four tea bushes was considered a single observation in the sample. The spectra were taken firstly from canopy scale, then from the fresh leaf stacks and finally from the dried and ground powders.

In order to maximize chemical variations, the sample comprised six different tea cultivars in the tea garden including Fuding dabai (FD), Fuyun 6# (FY), E cha 1# (EC), Tai cha 12# (TC), Huang dan (HD) and Mei zhan (MZ) (Table 1). These tea cultivars originated in different parts of China, and it has been reported that they vary greatly in foliar biochemical concentration (Astill et al. 2001; Lu et al. 1994).

4.2.3 Data collection

Spectral measurements of tea canopies

The spectral measurements of tea canopies were carried out in the tea garden between 10:30 and 13:00 on cloud-free sunny days (September, 2008). Canopy reflectance was measured using an ASD FieldSpec Pro FR spectrometer (Analytical Spectral Devices). The spectrometer covered a range from 350 to 2500 nm. The spectrometer covers a range from 350 to 2500 nm. The spectral resolution defined by Full-Width at Half-Maximum (FWHM) is: 3 nm at 700 nm, 10 nm at 1400 nm, and 12 nm at 2100 nm. The fiber optic was handheld approximately 10-20 cm above the top of the canopy, with a field of view of 25°, resulting in a spot area of approximately 15 to 60 cm^2 at canopy height. Less than 1% of the soil background was in the field of view due to the very

high leave area index (LAI) of the tea bushes. Six points were chosen for each tea bush and each spectral measurement at each point was the average of ten single measurements. The reflectance of a white spectralon panel (Labsphere, Inc.) was recorded between readings of the canopy to allow for normalization of light intensity. Four tea bushes together were considered to be one observation and the spectra from the four plants were averaged.

Spectral measurements of tea leaf stacks

After the measurements of the tea canopies, no less than 100 g of fresh leaves for each sample were clipped and taken to the lab (dark room) immediately for spectral measurements. When clipping, a single bud with three leaves from the top canopy of tea bushes was always selected. In order to keep the leaves fresh, the leaves were stored under cool temperatures in an ice box at 5 degrees Celsius and the time interval from collection to spectral measurement were less than two hours (Cho and Skidmore 2006).

The spectra of the fresh tea leaves were measured using the same ASD spectrometer (FieldSpec Pro FR with 25° fibre optics) as for the canopies, and the diameter of the spot size of the target leaves was 10cm. For each sample unit, piles of leaves, optically dense enough (10 layers +) to yield measurements of infinite reflectance (more than 200 leaves), were placed on top of a leaf tray made from black thick cardboard. A 50 W halogen lamp was positioned next to the sensor for the target illumination. The spectrum of each sample unit were recorded 9 times and then averaged. For each time, the leaves on the tray were randomly rearranged. To minimize the influence of BRDF (bidirectional reflectance distribution function), the tray was rotated horizontally (120 degree) after each third reading. For each recording, 10 measurements were taken and averaged. Before spectral measurement of leaf stacks, a white spectralon panel was measured for spectral standardization. After the reflectance of fresh leaves had been measured, the samples were dried and ground.

Spectral measurements of ground tea powder

The experiment was conducted in a dark laboratory. After drying the fresh leaves in an oven at 80°C over 10 hours, they were ground using an electric mill and passed through a sieve with a mesh width of 75 μm (ISO, 2005). For each sample unit, a 1 cm layer of the tea powder was placed in a dark container (size

of 10x10 cm) for spectral measurement. The reflective spectra were measured with the same ASD spectrometer described above with a 5° optic lens. The lens was fixed on a point about 12 cm above the black paper; therefore the diameter of the spot size on the measured target powder was 1 cm. The spectrum of each sample unit was recorded 3 times, and each record was the average of 10 measurements. A white spectralon panel was measured before for spectral standardization.

Wet chemistry analysis

The chemical analysis was carried out for dried and ground tea powders in the laboratory. Standard wet chemistry methods were used to determine the concentrations of total tea polyphenols, amino acids and soluble sugar. Concentrations of total tea polyphenols were determined using the Folin-Ciocalteu method at 765 nm (Ainsworth and Gillespie 2007); Briefly, 0.2 g of each sample was weighed, and 5 mL of 70% methanol at 70 °C was added (International Organization for Standardization (ISO) 14502-1). The extract was mixed and heated at 70 °C on a vortex for 10 min followed by cooling and centrifugation. One milliliter of the extract was diluted with water to 100 mL, from which 1.0 mL of the diluted sample extract was transferred in duplicate to separate tubes containing 5.0 mL of a 1/10 dilution of Folin-Ciocalteu's reagent in water. Then, 4.0 mL of a sodium carbonate solution (7.5% w/v) was added. The tubes were then allowed to stand at room temperature for 60 min before absorbance at 765 nm was measured against water. Total polyphenols was expressed as gallic acid equivalents in g/100 g material. The concentration of polyphenols in the samples was derived from a standard curve of gallic acid. The differences between parallel measurements were less than 3% for all chemicals**.**

Free amino acids were measured by ninhydrin colorimetry with spectrometry at 570 nm (Curran et al. 2001; Lee and Takahashi 1966). In a 25 mL volumetric flask, 1.0 mL of each tea infusion with 0.1 g of sample was added. This was followed by the addition of 0.5 mL of phosphate buffer solution (1/15 mol/L, pH 8.04) and 0.5 mL of 2% ninhydrin solution containing 0.8 mg/mL $SnCl2{\cdot}2H2O$. The mixtures in the volumetric flasks were then placed in a boiling water bath for 15 min. The probes were quickly cooled with cold water, and adjusted to 25 mL with water. After they were left standing for 10 min, the

absorbance values of the blue-purple products were measured against a reagent blank (replacement of tea infusion by an equal volume of water was used through the procedure). For other details refer to GB/T 8314-2002, a national standard method used to determine free amino acid content in tea infusions in China.

Soluble sugars were measured using anthrone colorimetry (Curran et al., 2001; Wen et al., 2005). Briefly, 0.1 g of dried sample (shattered, fineness: passage through 100-mesh) was weighed in a 10 mL centrifuge tube, to which 6–7 mL of 80% ethanol was added. The sample was heated in an 80℃ water bath for 30 min, then centrifuged (3000 r/min) for 5 min. The supernatant was collected, and the extraction was repeated twice (3000r/min for 10min each). The supernatant was collected into a flask, and 80% ethanol was added to a total volume of 50 mL. Then, 1 mL of the solution was taken, and 1.5 mL of water was added, followed by 6.5 mL of anthrone reagent. The sample was mixed and incubated at room temperature (18 – 30℃) for 15 min to allow colour development. The absorbance at 620 nm wavelength was then measured after the sample was cooled down. Table 4-1 shows the measured concentrations for each chemical component by cultivars.

Table 4-1 Descriptive statistics of the chemical variables measured in the laboratory (by cultivar)

	No. of observations	*Mean (mg g^{-1})*	*Variation coefficient**	*Minimum (mg g^{-1})*	*Maximum (mg g^{-1})*
		Tea polyphenols			
Fuding dabia	8	176.20	0.026	167.46	181.81
Fuyun 6#	8	180.67	0.041	172.21	195.35
E cha 1#	8	186.31	0.050	173.48	199.36
Tai cha 12#	8	208.72	0.022	203.90	213.77
Huang dan	8	270.92	0.035	260.63	288.99
Mei zhan	8	218.87	0.060	201.16	235.94
All	**48**	**206.95**	**0.163**	**167.47**	**288.99**
		Free amino acids			
Fuding dabia	8	23.86	0.012	23.46	24.17
Fuyun 6#	8	26.39	0.011	25.95	26.79
E cha 1#	8	29.40	0.066	27.08	32.57
Tai cha	8	24.73	0.011	24.42	25.12
Huang dan	8	21.12	0.022	20.55	21.93
Mei zhan	8	23.00	0.041	21.91	24.32
All	**48**	**24.75**	**0.113**	**20.55**	**32.57**
		Soluble sugars			
Fuding dabia	8	79.91	0.016	77.82	81.46
Fuyun 6#	8	84.86	0.023	81.32	87.72
E cha 1#	8	75.89	0.043	74.74	77.51
Tai cha 12#	8	75.19	0.029	73.00	78.76
Huang dan	8	71.35	0.017	69.17	72.92
Mei zhan	8	77.23	0.021	75.17	80.00
All	**48**	**77.41**	**0.058**	**69.17**	**87.72**

* Variation coefficient is the standard deviation divided by the mean value.

4.2.4 Data analysis

Division of training data and test data

The 48 observations in the sample were randomly divided into training data (N=30) and test data (N=18). The training dataset were used to calibrate the partial least squares regression model, and the performance of the model was validated by comparing the model predictions of the test data to the observations.

As the arbitrary division of training data and testing data may lead to possible biased results (Atzberger et al. 2003; Darvishzadeh et al. 2008c), the division procedure has been repeated for 100 times and the averaged r^2 and root mean square errors of the repetitions and their variances were calculated to evaluate the model performance.

Pre-processing of spectra

For the powder and leaf level analysis, the spectral bands between 350 nm and 400 nm and between 2300 nm to 2500 nm were excluded from the spectra because they were considered noisy (based on visual inspection). As a result, a total of 1900 wavebands from 401 to 2300 nm remained. For canopy level analysis, several band regions including 351 to 400 nm, 1351 to1420 nm, 1801 to 1970 nm and 2301 to 2500 nm displayed high levels of noise, and were excluded. The remaining 1660 wavebands were used for analysis. The spectral data excluding the noisy portions were mean-centered by subtracting their means before modeling.

Statistical analysis – partial least squares regression

Before the partial least squares regression (PLSR) analyses were applied to a training set, the data is mean centered. Mean centering has the effect of enhance the subtle differences between the spectra (Workman and Springsteen, 1998) and processed using Matlab 7.1. The performances of the PLSR models were assessed by the coefficient of determination (r^2) between predicted and measured concentrations and the root mean square error of prediction (RMSEP, Eq. 4-1) on test dataset (Cho et al. 2007). To determine the factor number of the partial least squares regression models, leave-one-out cross validation was used (Cho et al. 2007; Schlerf et al. 2005). In order to avoid the collinearity problem and to maintain the model simplicity, the criterion to add an extra factor to the partial least squares regression model was that it had to reduce the root mean square error of cross validation (RMSECV) by less than 2% (Kooistra et al. 2004).

$$\text{RMSEP}=\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \qquad \text{Eq. 4-1}$$

Where n is the number of test data, y_i is the observed value of data point i and $\hat{y}_i$ is the estimated value based on the model calibrated by training data.

The calibration equation coefficients (*b*-coefficients) were used to determine the importance of spectral bands in PLSR model calibrations (Haaland and Thomas 1988; Wold et al. 2001). In the model, *b*-coefficients represent the contribution of each predictor (waveband) to the model. As Gomez et al. (2008) suggested, a threshold of one standard deviation (±σ) for *b*-coefficients were used, and a wavebands was considered significant if the corresponding *b*-coefficients (absolute value) was larger than one standard deviation. As the model was run for 100 times with different training data, the bands that were identified as significant for more than 90 times (90%) were regarded as the robust and important ones.

4.3 Results

Figure 4-1 shows the averaged spectra and the standard deviations for 48 observations for the powder, leaf and canopy level respectively. For powders, variability peaked along the red edge and was at a minimum at wavelengths shorter than 500 nm and between 1400 and 2000 nm. For fresh leaves, variability was low for the visible bands and the two main absorption peaks of water (1400 nm and 1950 nm). For canopies, variability was generally higher at near infrared plateau and lower for the visible regions.

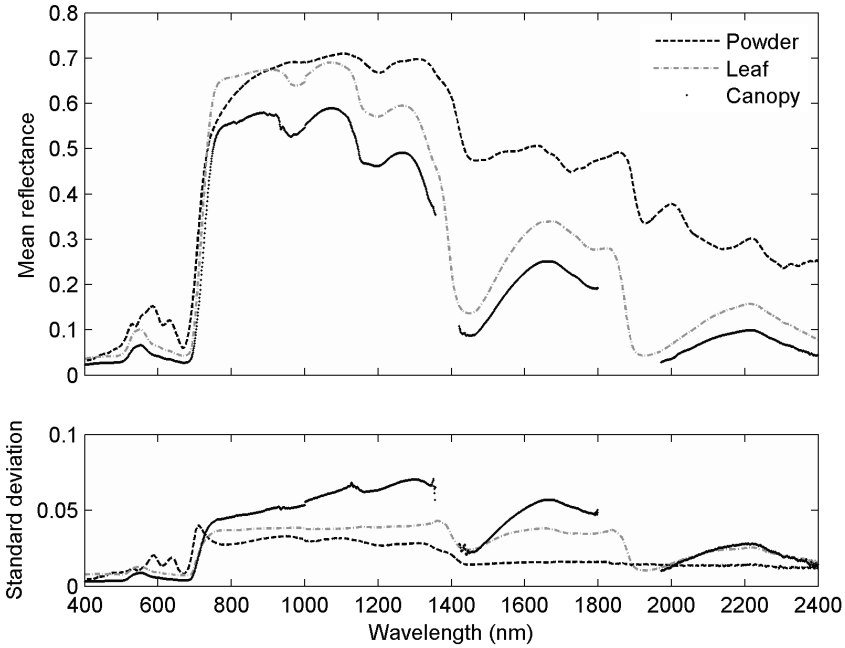

Figure 4-1 Mean reflectance (top) and standard deviation of reflectance (bottom image) of dried powders, fresh tea leaves and tea canopies (n=48).

Observed versus predicted concentrations of tea polyphenols, free amino acids and soluble sugars for both training (N=30) and test (N=18) data are shown in Figure 4-2. This figure illustrated the result of one of the 100 model iterations. Comparing the results of three different levels, total tea polyphenols, free amino acids and soluble sugars were estimated with the highest r^2-values and lowest RMSE values at powder level for both the calibration and validation data sets. The lowest accuracy of prediction was obtained at leaf level. For total tea polyphenols and free amino acids, high r^2 values were obtained at all three levels, while retrieving soluble sugars from leaf and canopy levels resulted in lower r^2 (less than 0.6) and a larger RMSE.

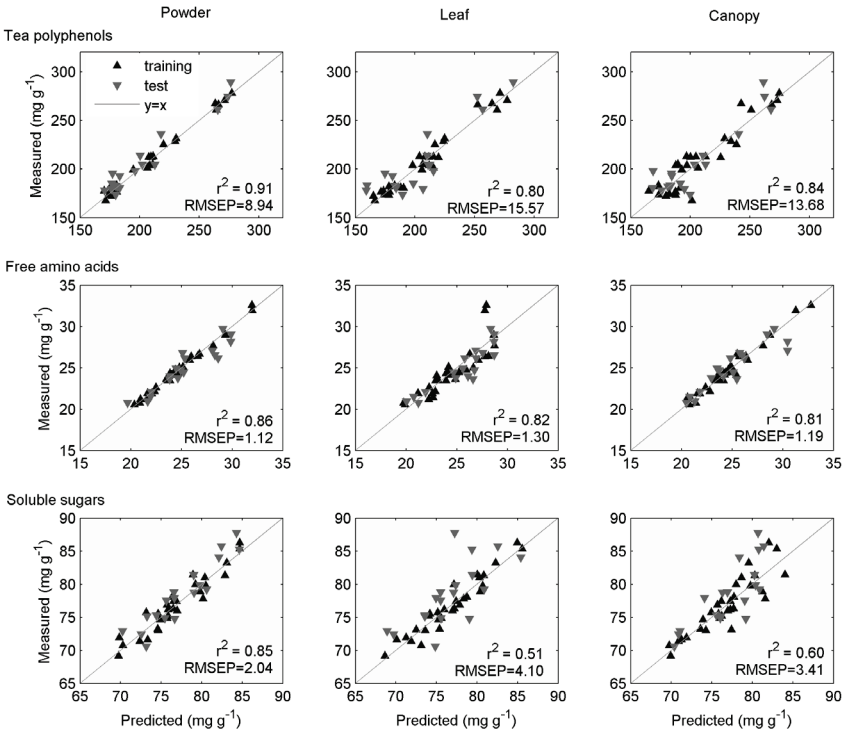

Figure 4-2 Exemplary scatter plots describing the measured and predicted foliar biochemicals (total tea polyphenols, free amino acids and soluble sugar) for training and test using powder, leaf and canopy spectra (mean centered). r^2 is coefficient of determination between model predictions and measured chemical concentrations on test data set, and RMSEP is the root mean square error of test data prediction.

Table 4-2 shows the number of PLS factors, the model errors and the error variations in the 100 iterations. Small error variations were observed as an indication of the model robustness. On test data set, the overall accuracy of the partial least squares regression models was good in terms of the averaged prediction errors, which were below 10% of the mean values of the chemicals (relative RMSEP) for all three levels. For each chemical component prediction, the highest accuracy was achieved for tea powders and the moderate accuracy was achieved for tea plants.

Table 4-2 Performance of partial least squares regression for predicting total tea polyphenols, free amino acids and soluble sugar

	PLS factors No.	r^2 (test) ± St.Dev.	RMSEP * ± St.Dev. (mg g^{-1})	Relative RMSEP ** ± St.Dev. (%)
Total tea polyphenols				
Powder	2	0.89±0.04	11.31±1.69	5.47±0.80
Leaf	4	0.81±0.09	14.62±3.07	7.09±1.47
Canopy	3	0.83±0.09	13.99±3.06	6.79±1.54
Free amino acids				
Powder	4	0.81±0.06	1.39±0.22	5.50±0.09
Leaf	3	0.76±0.11	1.43±0.31	5.74±1.22
Canopy	4	0.77±0.07	1.41±0.25	5.73±1.01
Soluble sugars				
Powder	3	0.78±0.11	2.13±0.43	2.75±0.55
Leaf	2	0.46±0.16	3.43±0.58	4.46±0.73
Canopy	3	0.56±0.13	3.12±0.55	4.03±0.72

* RMSEP is root mean square error of prediction on test dataset; ** Relative RMSEP is RMSEP divided by mean value.

The important wavelength selection processing was performed for the three chemicals of interest at the level of powder, leaf and canopy separately. Figure 4-3 shows the frequency of each waveband identified as important from the 100 iterations.

Table 4-3 gives an overview of the most important wavelength regions (the bands which were identified important more than 90 times) and the chemical bonds they have been identified with. References are given to the readers in order to explore the related chemical bonds and their absorption features (Table 4-3). Most wavelengths selected were not exactly at the known absorption features but within ± 12 nm of them (followed by the 12 nm range set up in (Curran et al., 2001)).. The wavebands causally correlated (± 12 nm) with the absorption features of the target chemicals could be found important at both powder and leaf scales, but had less chance to be identified at the canopy scale.

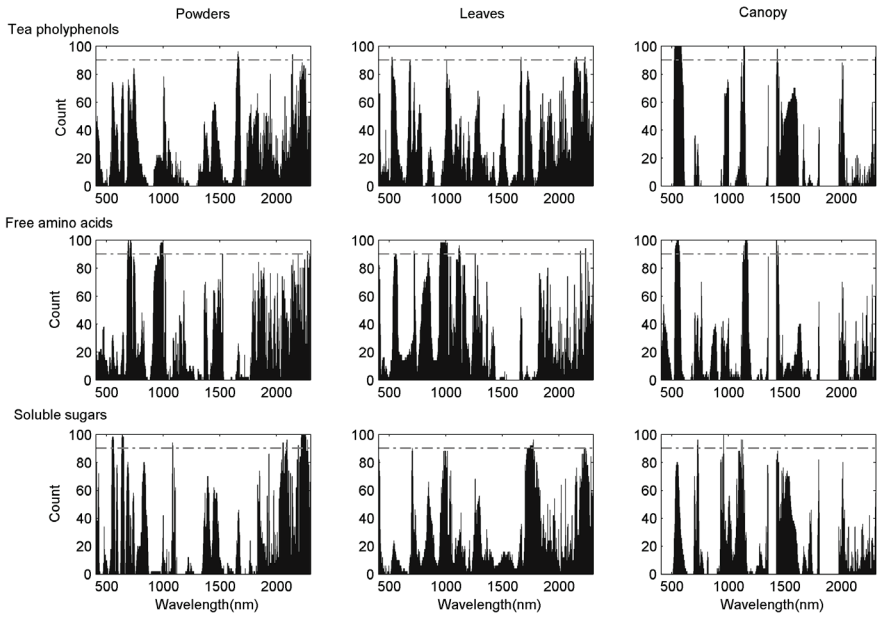

Figure 4-3 Bar graphs showing the count numbers of each wavebands identified as the important ones, which are selected from the partial least squares regression models iterated for 100 times. The dashed line highlights the confines of 90 times.

Table 4-3 The important wavelengths for prediction of tea polyphenols, free amino acids and soluble sugars, their possible chemical bonds and source of references (the wavebands in bold are causally related to the biochemicals of interest; U is unattributable to a biochemical absorption feature)

		Wavelengths found important	**Attribute to (source)**
Total tea polyphenols	Powder	**1656-1660**	Phenolic, catechin [a,b]
		1663-1667	U
		2139-2140	N-H stretch[c,d]
	Leaf	522-528	Colour information (green) [e]
		680-683	Chlorophyll, nitrogen [f]
		1006-1007	O-H stretch, 2nd overtone [c]
		1658-1663	Phenolic, catechin [a,b]
		2134-2135	N-H stretch [d]
		2146-2152	U
		2226-2228	C-H stretch [c]
	Canopy	520-585	Chlorophyll, nitrogen [g]
		1133-1140	C-H stretch, 2nd overtone [c]
		1429-1430	C-H stretch, C-H deformation [c]
		2009-2010	O-H deformation [c]
		2295	N-H stretch, N= O stretch [c]
Free amino acids	Powder	684-697	Chlorophyll, nitrogen [d]
		706-718	Chlorophyll, nitrogen [d]
		969-997	O-H bend, 1st overtone [c]

			O-H stretch, 2nd overtone [c]
		1015	N-H stretch (Protein) [c]
		1520-1523	Amino acids [h]
		2188	N-H bend, 2nd overtone (Protein, Nitrogen) [c]
		2271-2272	C-H stretch, O-H stretch [c]
		2291-2293	N-H stretch, C = O stretch (Protein, nitrogen) [c]
	Leaf	553-560	Chlorophyll, nitrogen [g]
		716-723	Chlorophyll, nitrogen [d]
		947-999	O-H bend, 1st overtone [c]
		1005-1012	N-H stretch (Protein) [c]
		1117-1124	C-H stretch, 2nd overtone [c]
		1257	U
		2186-2188	N-H bend, 2nd overtone (Protein, Nitrogen) [c]
		2231-2232	C-H stretch (Protein) [c,i]
	Canopy	529-566	Chlorophyll, nitrogen [d,g]
		1131-1136	C-H stretch, 2nd overtone [c]
		1146-1165	U
		1421-1424	C-H stretch, C-H deformation [c]
		1435-1436	U
Soluble sugar	Powder	546-561	Chlorophyll [d,g]
		634-648	Chlorophyll [c]
		1082-1083	U
		2054-2058	N=H bend, 2nd overtone [c]
		2086-2091	Sugar, starch [c]
		2189-2190	N-H bend, 2nd overtone [c]
		2217-2252	C-H stretch/ O-H stretch , O-H deformation [c]
		2269	Cellulose, sugar [c]
	leaf	701-703	Chlorophyll, nitrogen [d]
		1717-1718	U
		1720-1760	U
		1768-1771	Cellulose, sugar [c]
		2220-2234	C-H stretch [c]
	Canopy	724-729	Chlorophyll, nitrogen [d]
		954-955	U
		1097	U
		1111-1124	C-H stretch, 2nd overtone [c,i]

a. (Lu et al. 2005; Soukupova et al. 2002)
b. (Soukupova et al. 2002)
c. (Curran 1989)
d. (Mutanga et al. 2004b)
e. (Workman and Springsteen 1998)
f. (Cho and Skidmore 2006)
g. (Penuelas et al. 1994)
h. (Curran et al. 2001)
i. (Kumar et al., 2001)

4.4 Discussion

We explored the differences in spectral variability among three different measurement levels of tea plants. The variability in spectra reflectance of dried tea powders is mainly correlated with the amount of chemical compounds, because absorption by a dry powder is unaffected by water as well as effects caused by leaf cell structure (Curran et al. 2001; Kokaly and Clark 1999). At the leaf level, reflectance variability (401nm-2400nm) is mainly a function of leaf structure, biochemical concentrations and water content (Darvishzadeh et al. 2008a). At the canopy level, reflectance variability encompasses additional factors such as LAI (leaf area index) and canopy architecture (Darvishzadeh et al. 2008a; Gitelson et al. 2003).

Comparing the results of the three scales, the highest accuracy model was achieved at the powder scale (Table 4-2). In previous studies, some phenolic substances of tea, including total tea polyphenols, catechins, epigallocatechin gallate (EGCG) and epicatechin (EC) have been successfully estimated using near infrared spectra for dried tea powders and dried leaves (Chen et al. 2008a; Luypaert et al. 2003). Free amino acids and soluble sugars in tea have also been estimated using reflectance spectroscopy at the powder and leaf levels (Bian et al. 2010); however, these studies lacked systematic comparison among different levels and cultivars.

In this study, the lowest prediction accuracies were obtained at the fresh leaf scale. The predictions for total tea polyphenols and free amino acids are more accurate than that of soluble sugars. Bian et al. (2010) have reported that for one cultivar of green tea, the tea quality related total tea polyphenols and free amino acids could be retrieved from fresh leaf stacks using reflectance spectroscopy under controlled laboratory conditions. The decrease in predictive ability from powder to leaf level was expected, because many factors, such as water content, leaf angle and leaf cell structure could influence the leaf spectra and mask the minor absorption features (Figure 4-1) (Curran et al. 1992).

The prediction results at the canopy level slightly outperformed the results from the leaf level (Table 4-2). This finding is supported by Asner and Martin (2008) who found that at canopy level, partial least squares regressions at constant LAI levels and homogenous canopy structure are equally or more robust than those

made at the leaf level. They found that in highly foliated canopies (LAI>5), multiple scattering in the upper canopy leaf layers might enhances the expression of leaf properties, including chlorophyll and foliar nutrient (nitrogen, phosphors). However, at the canopy level, the influential wavelength channels that were correlated with biochemicals of interest (Table 4-3) have less chance to be identified important, compared to powder and leaf scales. This demonstrated poor signal propagation from powder to the leaf scale, and from the leaf to the canopy scale (Verhoef 1984; Yoder and Pettigrew-Crosby 1995).

Compared to total tea polyphenols and free amino acids, predicting soluble sugars using partial least squares regression model had a lower accuracy, especially at the leaf and canopy scale. The low predictive power may be attributed to the narrow range of sugars concentrations in this study (Table 4-1). The content of soluble sugar has the lowest coefficient of variation (0.06), which is much smaller compared to that of the other two chemicals. In addition, the content of soluble sugars was lower than 10% in the dry matters of tea leaves, which also increased the prediction difficulty.

The important bands identified in this study match wavelengths described by previous studies within 12 nm distance of the known absorption features (Table 4-3) (Curran 1989; Curran et al. 2001), probably due to instrument calibration. In general, this spectral reflectance signal suffers from sensor drift, caused by temperature change, aging of electronic components, and etc. (Ferwerda et al. 2006a). It will result in a slight mismatch between the labels and the spectral position of the actual reflectance. Some of the important bands from powder and leaf spectra can be linked to the known absorption features of the chemicals of interest. It was reported that one of the absorption features of catechins, which is the main substance of tea polyphenols, is found at 1645 nm (Lu et al. 2005). Soukupova et al. (2002) described an absorption feature related to phenolic at 1648 nm. And we found that the wavebands from 1656 to 1660 nm at powder level and wavebands from 1658 to 1663 nm at leaf level played an important role in predicting tea polyphenols; for amino acids, we found 1520-1523 nm contributes directly to the free amino acids estimation from powder spectra, for it has been stated the absorption feature related to amino acids at 1510 nm (Curran et al. 2001). For soluble sugars, we found the influential bands located near 2086 nm and 2242nm at the powder level and 1768-1771nm at the

leaf level, which is coincident with the absorption of sugar and starch (Curran 1989). The other influential wave bands identified by partial least square regression are indirectly related to the chemicals of interest. For example, amino acids and protein is highly correlated, which could explain that when predicting free amino acids, some of the influential bands are close to the protein absorption features around 1020 nm, 2180 nm and 2240 nm (Curran 1989). We speculate, if the impact of foliar mesophyll cell structure or canopy structure on the reflectance spectra overshadows the absorption peaks of foliar chemicals, the partial least squares regression will select surrogate spectral features that relate to other chemicals but have strong correlation with the one intended to be estimated.

4.5 Conclusions

The following conclusions were drawn from this study:

(1) Our results suggest that biochemical components (total tea polyphenols, free amino acids and soluble sugars) of tea quality can be quantitatively estimated from powder, leaf and canopy spectroscopy. It may have the potential to predict foliar biochemical compositions in tea from space. However, only after atmospheric effects and soil background effects are eliminated, can we estimate quantitatively the foliar parameters related to tea quality. At this stage, we will recommend based on our result at canopy scale, the bands centered at 552 nm, 548 nm, 726 nm, 955 nm, 1097 nm, 1117 nm, 1133 nm, 1137 nm, 1155 nm, 1422 nm, 1430 nm, 1436 nm, 2010 nm and 2295 nm for the satellite based remote sensing for tea quality monitoring. These bands needs further tests in satellite based remote sensing for tea quality estimation at large scale.

(2) Predicting soluble sugars from leaf and canopy spectra was achieved with lower accuracy, likely due to the low concentrations and the small variation of soluble sugars in tea leaves.

(3) The most important bands selected by partial least square regression are not consistent across the three measurement scales. Compared to the powder

and leaf scales, more indirectly related wavelengths were found at the canopy level.

In summary, the prediction accuracies achieved at canopy level in this study gives rise to positive expectations from airborne and space-borne sensors for repeated mapping of tea quality, at landscape or regional scales.

CHAPTER 5

TEA-PROSPECT: A LEAF OPTICAL MODEL FOR TEA INCORPORATING POLYPHENOLS

This chapter is based on: Meng Bian, Andrew K. Skidmore, Tiejun Wang, Yanfang Liu & Teng Fei. Tea-PROSPECT: a leaf optical model for tea incorporating polyphenols. *International Journal of Applied Earth Observation and Geoinformation* (in review).

Abstract

A modified PROSPECT leaf optical model for tea, incorporating total tea polyphenols being indicators of tea quality, was developed. Foliar hemispherical reflectance was acquired from 400 nm to 2500 nm using an ASD spectrometer, and compared with the simulated spectra using the original and the improved leaf models. The results showed that the new model was more accurate when simulating spectra. During inversion, the retrieval of total tea polyphenols using the modified PROSPECT model resulted in an r^2 of 0.52 with a relative RMSE (RMSE/mean) of 0.23 for fresh tea leaves at different phenological stages; an r^2 of 0.66 with a relative RMSE of 0.10 for fresh tea leaves of different tea cultivars; and an r^2 of 0.69 with a relative RMSE of 0.17 for the two datasets pooled together. To test the retrieval of the physically based model, a reported empirical regression model was also applied to the same datasets. The improved PROSPECT tea model retrieves the concentration of tea polyphenols from the spectra for any growth stage and any tea cultivar with moderate accuracy, thereby demonstrating generality and robustness in its ability to model across a wide range of conditions.

5.1 Introduction

Hyperspectral remote sensing allows (virtually) continuous radiometric measurements from which the biochemical characteristics of vegetation may be derived with high accuracy (Curran 1989; Knox et al. 2011). In general, there are two different approaches to estimate vegetation biochemical parameters from hyperspectral data: the first one is the empirical approach, in which statistical correlations between leaf reflectance (or some vegetation indices) and leaf biochemical content have been established (Cho and Skidmore 2006); the second one is based on physical models, using the physical laws of electromagnetic radiation within leaves (Darvishzadeh et al. 2008b; Jacquemoud et al. 1996). Although the statistical relationship approach is relatively easy to set up, prediction results have been recognized as being sensor-specific and dependent on site and sampling conditions (Darvishzadeh et al. 2008b). In contrast, the physically based models are more robust and portable, and the development of such models has resulted in better understanding of the interaction between light and plant leaves (Jacquemoud et al. 2000). However, the physically based models remain models, that is, simplifications of reality, and require empirical data to weight the model parameters.

Tea (*Camellia sinensis* L) is a major economic crop. The quality of tea, and hence its value, is mainly reflected in its smell and taste, which can be determined from the composition of fresh leaves (Atoui et al. 2005; Liang et al. 2003; Yamamoto et al. 1997). Accurate and quantitative estimation of leaf biochemicals is useful for predicting tea quality before plucking, and improving tea garden management (Bian et al. 2010). Tea polyphenols, accounting for 18-36% in weight of the total dry matter of tea leaves, have been generally regarded as the most important biochemical components of tea quality, for their contribution to the special flavour of tea and possible beneficial effect on human health (Chen et al. 2008a; McKay and Blumberg 2002). Instead of time-consuming wet chemistry, laboratory-based near infrared spectroscopy (NIRS) has been developed (Bureau et al. 2009; Pojić et al. 2012; Sáiz-Abajo et al. 2006). It has been used as a fast and non-destructive way to predict alkaloid and phenol substances in tea (dried tea samples) since the 1990s (Schulz et al. 1999). With the development of hyperspectral remote sensing, the research has been extended to measure tea quality-related biochemicals on a larger scale (fresh

leaves and tea canopies) by field spectrometer or air/space-borne sensors. For example, both the studies by Bian et al. (2010) and Dutta et al. (2011) show the potential of hyperspectral techniques for estimating *in situ* tea quality. However, they used either multivariate regression or vegetation indices, both belonging to the empirical, or statistical, approach, resulting in derived correlations that may change in space and time (Darvishzadeh et al. 2008b). Therefore, it is of interest to develop deductive or physically based models to retrieve the biochemical parameters as quality indicators of tea using hyperspectral data.

Among physically based models, radiative transfer models have been considered simple and effective for calculating plant reflectance, and through inversion to estimate vegetation biochemical parameters (Jacquemoud et al. 2000; Verhoef and Bach 2007). The optical properties of a plant canopy largely depend on the optical properties of the leaves and the soil background. Thus, it is critical to understand physiological process of leaves and develop leaf optical properties models (Botha et al. 2010). The leaf radiative transfer model PROSPECT has been validated and applied in numerous studies (Feret et al. 2008; Fourty et al. 1996; Jacquemoud et al. 2009). The original version of the PROSPECT model describes reflectance from and transmittance within a fresh leaf in the 400-2500 nm spectral range as a function of a leaf mesophyll parameter (N), a spectral refractive index (n), plus chlorophyll and water concentrations and their corresponding spectral absorption coefficients (Jacquemoud and Baret 1990). In later years, the model was refined to include the influence of other biochemical components in the leaf, such as protein, cellulose, lignin, and starch. Both the original and updated versions of the model may be inverted (Botha et al. 2006; Jacquemoud et al. 2000). Many studies have been conducted to estimate and quantify vegetation biochemical variables such as chlorophyll using the PROSPECT leaf model alone or in combination with a canopy reflectance model during the past two decades (Houborg and Boegh 2008; Jacquemoud et al. 1996; Le Maire et al. 2004; Miao et al. 2011). The research covers agricultural crops and forests with differing retrieval accuracies. The current version of the PROSPECT model may not be suitable for the prediction of biochemical concentrations as quality indicators for tea, because this model ignores the contribution of tea polyphenol compounds to the spectral reflectance. Polyphenols have not been regarded as

playing a major role in the composition of dry leaf matter for most plants (Jacquemoud et al. 1996), but they do have a high concentration in tea leaves.

In this study, we develop an improved PROSPECT leaf optical model for tea that incorporates polyphenols. The improvement is based on an updated version of the PROSPECT model (Baret and Fourty 1997; Jacquemoud et al. 2000), in which total dry matter content is considered, instead of individual measures of protein, cellulose, lignin, and so on. The specific objectives of the study are: (*a*) adding tea polyphenols as a parameter to the PROSPECT model to improve its performance for predicting the spectra of tea leaves; (*b*) validating the new model using a forward process (from biochemical parameters to spectrum) and by inversion (from spectrum to bio-chemical parameters); (*c*) comparing the retrieval accuracies resulting from the physical and the statistical approach reported in previous studies.

5.2 Materials and Methods

5.2.1 Study area and sampling strategies

Tea leaves were collected from an experimental field (latitude 30°28'41"N, longitude 114°21'48"E) near Wuhan, which is located in the middle-lower Yangtze plain in China. The climate is suitable for several tea cultivars, with an average annual precipitation of more than 1000 millimeters as well as 1950 hours of sunshine.

Field work was carried out in early June and also at the end of September, 2010. In the first campaign, leaf samples from the tea of 'Fuding dabai' were collected. This tea cultivar is widely cultivated in China and produces good quality green and white teas (Yao et al. 2008). When building a physical model, one difficulty is that absorption features of dry matter may be hidden by predominant absorption features of water. An effective solution to overcome this is to collect datasets containing leaves with a wide distribution of biochemical concentrations. As previous research has shown that tea polyphenol concentrations varied a lot between young and old tea leaves (Yamamoto et al. 1997), in this campaign, the leaves of the 'Fuding dabai' tea cultivar were plucked at different phenological stages. Using a single season's growth, tea leaves were collected, starting at the first leaf through to the seventh leaf of a branch. For each of the seven phenological stages, we picked the leaves

randomly in the tea garden until the weight reached the requirement of wet chemical analysis (more than 70 g), with 4 duplicate samples being collected to gain more representative values (Houborg and Boegh 2008).

To test the physical model on different tea cultivars, another field survey was carried out in late September. Six different tea cultivars in the tea garden, namely Fuding dabai (FD), Fuyun 6# (FY), E cha 1# (EC), Tai cha 12# (TC), Huang dan (HD), and Mei zhan (MZ), were selected for this study. These tea cultivars originated in different parts of China, and it has been reported that they vary greatly in foliar biochemical concentration (Astill et al. 2001; Lu et al. 1994). We picked tea leaves from the top canopy of tea bushes. For each of the six tea cultivars 4 duplicate samples were collected, ensuring, as before, that enough leaves were collected for the wet chemistry assay.

Thus, two sub-sets of data were collected. One sub-set represented tea leaves under different growing stages (n = **7**x4); the other was collected from six different cultivars representing tea leaves in general (n=**6**x4). For each growing stage and for each tea cultivar, four samples were collected and their measurements averaged to obtain more general characteristics of the tea. In order to minimize wilting, all the picked leaves were stored at 5 degrees Celsius and the spectral reflectance measured within two hours of picking (Cho and Skidmore 2006).

5.2.2 Data collection

Leaf spectral measurements

Hemispherical reflectance measurements were taken with a portable ASD FieldSpec Pro FR spectrometer (Analytical Spectral Devices) equipped with a RTS-3ZC integrating sphere coated with Zenith diffuse polymer material interiorly (ASD Inc.) (Stuckens et al. 2009). For each sample unit, 10 to 20 leaves were randomly selected for the spectral measurements, and the averaged reflectance of the selected leaves was calculated to represent the sample. The abaxial (upper faces) reflectance of each leaf was measured with a 3 nm spectral resolution in the 350-1000 nm range, and a 10 nm resolution in the 1000-2500 nm range. Data were corrected for the reflectance of a nominal 99% Zenith® diffuse reference standard of the integrating sphere (Delalieux et al. 2008).

The reflected spectra were originally recorded with a re-sampling interval of 1 nm. To suppress noise in the data, a moving Savitzky-Golay filter with a window size of 7 nm was applied to the averaged reflectance measurements (Workman and Springsteen 1998). Then, the spectra were re-sampled every 5 nm to match the wavelength interval of the PROSPECT model used in this study (Jacquemoud et al. 2000). Bands with a wavelength of less than 400 nm displayed high levels of noise and were excluded from our study. Thus, for each leaf sample unit, the resulting 421 wavebands covered the 400-2500 nm range, at 5 nm wavelength intervals.

Specific absorption coefficients for tea polyphenols

When determining the optical constants of leaf materials for the PROSPECT model, an important step is to estimate the specific absorption characteristics of leaf components. The specific absorption coefficient describes absorbance of light per unit path length (usually in centimeters) and per unit of mass concentration (usually grams or micrograms per liter) (Feret et al. 2008). Based on the specific absorption coefficients and the chemical concentrations, the absorption features of a specific leaf material can then be calculated using the following equation (Eq. 5-1).

$$k(\lambda) = \sum_i k_{spe,i}(\lambda) \times \frac{C_i}{N} \qquad \text{Eq. 5-1}$$

Where λ is the wavelength, N is the structure parameter of the leaf, C_i is the concentration of component i, and $K_{spe,i}$ is the corresponding specific absorption coefficient.

We introduced a new parameter, namely tea polyphenols, to the PROSPECT leaf optical property model. The absorption coefficient of polyphenols extracted from tea leaves in a solvent of pure acetone (with a concentration measured in units of grams per liter) was measured using a UV-visible spectrophotometer (SHIMADZU UV-1601PC, from 400-1000 nm) and a Fourier Transform Near-Infrared Spectrometer (FT-NIR Antaris II) in transmittance mode (from 1000-2500 nm). The pure acetone solvent was simultaneously measured to eliminate its influence on the absorption features of tea polyphenols. Spectra were originally scanned in 1 nm steps, and then the wavelength interval was re-sampled to 5 nm for model calculation.

In our study, instead of using the strategy adopted by Jacquemoud et al. (Jacquemoud and Baret 1990; Jacquemoud et al. 1996), who deduced the specific absorption coefficients of leaf biochemical components using *in vivo* spectral measurements and measured concentrations, perhaps taking covariation among some of the biochemicals into account, we measured the specific absorption coefficient of tea polyphenols, to investigate whether and how tea polyphenols, as a separate parameter, influence the spectral variation in tea leaves.

Measurements of biophysical and biochemical parameters

After the spectral measurements, we measured the weight, the blade thickness, and the specific leaf area (SLA, usually square centimeters per unit leaf dry weight) for each observation in the leaf samples. Then, the leaves for each observation were divided into two piles for biochemical analysis in the laboratory. Pile I was used for chlorophyll content measurement. According to the international standard, the tea leaves were homogenized in acetone prior to the spectroscopic determination (at 645 nm and 663 nm) of chlorophyll *a* and *b*, from which the total chlorophyll concentration is estimated (Curran et al. 2001). Pile II was used to calculate the dry matter content and determine the concentrations of total polyphenols. For this the tea leaves were dehydrated and then ground using an electric mill. According to the international standard for the determination of biochemical contents in tea, total tea polyphenols were determined by the Folin-Ciocalteu colorimetry method and spectrometry at 765nm (ISO 14502-1:2005). Through estimating the weight change from fresh leaves to dried leaves, the water content for each observation was derived.

As required by the PROSPECT model, the measured water content and biochemical concentrations were expressed in mass per unit leaf area using the specific leaf area.

5.2.3 Improved PROSPECT model for tea (tea-PROSPECT)

Being a radiative transfer model, PROSPECT assumes that a leaf is a stack of N (a parameter charactering the leaf mesophyll structure) identical elementary layers separated by N–1 air spaces, and calculates the leaf hemispherical reflectance and transmittance from 400 nm to 2500 nm (Jacquemoud et al. 2009). In this model, scattering is described by the refractive index of leaf

materials (n) and by the mesophyll structure parameter N; in the original version absorption was modeled using only the absorption characteristics of total chlorophyll and water. The PROSPECT model was later refined to introduce new leaf biochemical components such as protein, cellulose, lignin and so on. For simplicity reasons, Baret and Fourty (1997) and Jacquemoud et al. (2000) considered the dry matter content as a whole instead of considering those biochemical compounds individually. This version of the PROSPECT model has been widely used (Darvishzadeh et al. 2008b), and our improvements for the tea-PROSPECT model are based on this latter version.

The parameters for PROSPECT are derived from experimental datasets, for example, the LOPEX experiment from 1993. Although many types of vegetation were covered by those datasets (Feret et al. 2008), the lack of specificity in the parameters, such as the arbitrary value for absorption characteristics of dry matter, makes it less accurate for a specific plant species. In tea leaves, total tea polyphenols make up more than 20 percent of the dry mass, while in most other plants, polyphenols, belonging to the secondary metabolites, exist in low concentrations. Therefore, when modeling the reflectance/transmittance spectrum of tea leaves, tea polyphenols clearly need to be taken into account to ensure an accurate prediction.

The improved PROSPECT model for tea (tea-PROSPECT) calculates the tea leaf hemispherical reflectance as a function of five input parameters, i.e., the leaf structural parameter, N (unitless); the leaf chlorophyll *a*+*b* concentration, C_{a+b} (μg cm^{-2}); the equivalent water thickness, C_w (g cm^{-2}); the concentration of total tea polyphenols, C_{tp} (g cm^{-2}); and the dry matter content excluding tea polyphenols, C_m (g cm^{-2}). Tea polyphenols have been separated off as a specific input parameter of the model, in an attempt to model the leaf spectra of tea with higher accuracy and explain the absorption features that cannot be attributed to water or to other biochemicals.

The parameter N relates to the cellular arrangement within the leaf. Jacquemoud and Baret (1990) indicated that this parameter is positively correlated with the specific leaf area (SLA) and proposed a hyperbolic relationship to quantify N. In this study, instead of using empirical correlations established using limited datasets, we determined N through the physically based model by minimizing

the difference between simulated and measured spectra data, with a reasonable range of N (Ceccato et al. 2001). Since the influence of pigments, water and dry matter (mostly composed of cell wall molecules such as cellulose and lignin) on the visible and near infrared spectral regions was well modeled by PROSPECT, the specific absorption coefficients for chlorophyll *a* and *b* (K_{a+b}), water (K_w) and dry matter (without tea polyphenols, K_m) were assumed to also be applicable to tea-PROSPECT. For the newly introduced polyphenol parameter, the specific absorption coefficient (K_{tp}) was determined for tea polyphenols in the 400-2500 nm region.

5.2.4 Spectral simulation and model inversion

The performance of the tea-PROSPECT model was firstly evaluated against the previous PROSPECT model in the forward phase. Then the improved model was inverted to test whether the tea polyphenol total is retrievable as quality indicator of tea by using the physically based approach. Two datasets – tea leaf samples collected from one tea cultivar, as well as from multiple tea cultivars – were used for model validation, for both of the forward and the inversion process. All results were based on the mean value obtained from four duplicate field samples.

Simulated reflectance vs. measured reflectance

We first tested our improved model in direct mode (forward process), by simulating the reflectance spectra of fresh tea leaves using seven different growth stages of one tea cultivar (Fuding dabai), as well as using six different tea cultivars from 400-2500 nm. The measured concentrations of pigments, water, total tea polyphenols and dry matter, as well as the estimated values of the mesophyll structure parameter N, were used. The error between the simulated reflectance and measured reflectance using PROSPECT and tea-PROSPECT were calculated separately using the root mean square error (RMSE) – see below for details.

Retrieving total tea polyphenols

We inverted the improved model to retrieve the concentration of total tea polyphenols, as a quality indicator of tea. A widely used method to invert a radiative transfer model is by look-up table (LUT). It has been applied as an effective alternative to numerical optimisation and neural network methods, as

it offers a global search while showing less unexpected behavior (Combal et al. 2002; Schlerf and Atzberger 2006; Tang et al. 2006). Due to the multiple input variables of the inverted model (the mesophyll structure parameter N, total chlorophyll, water, tea polyphenols, and dry matter), it is possible that more than one set of solutions could lead to a similar match between the measured and simulated spectra (Botha et al. 2006), resulting in improper estimation of the leaf biochemical constituents. To achieve high accuracies for the estimated parameters, the dimension of the look-up table must be sufficiently large.

In our study, the model inversion was performed in Matlab (The MathWorks Inc., 2011), and a look-up table was built in advance of the actual inversion. In the forward calculation of the tea-PROSPECT model 100,000 parameter combinations were randomly generated and applied. The initial model parameters were randomly selected from uniform distributions. To prevent too-wide parameter spaces and to reduce the calculation time, the range of parameters was fixed, based on a literature review as well as on prior knowledge relating to the physiology characteristics of tea (Combal et al. 2003; Darvishzadeh et al. 2008b) (Table 5-1). By comparing the root mean square error between the simulated (by LUT) and measured spectra, the optimal combination of parameters corresponding to the lowest error values for a given spectra can be selected.

Table 5-1 Specific range of five input parameters of the improved PROSPECT model.

Parameter	*Unit*	*Min*	*Max*	*Reference*
Leaf structure parameter (N)	None	1	3.5	(Jacquemoud and Baret 1990)
Leaf chlorophyll content (C_{a+b})*	μg cm^{-2}	15	100	—
Equivalent water thickness (C_w)	g cm^{-2}	0.01	0.02	(Darvishzadeh et al. 2008b).
Concentrations of total tea polyphenols(C_{tp})*	g cm^{-2}	0.001	0.003	(Yamamoto et al. 1997).
Dry matter content (without tea polyphenols) (C_m)	g cm^{-2}	0.005	0.01	(Darvishzadeh et al. 2008b).

* The minimum and maximum values are based on knowledge from the field

Retrieving total tea polyphenols for tea -- accuracy assessment

The retrieval accuracies were evaluated using r^2, RMSE, as well as the relative RMSE (RMSE/mean) of the predicted and observed total tea polyphenol content (Eq. 5-2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (TP_{measured} - TP_{estimated})^2}{n}} \quad \text{Eq. 5-2}$$

Where $TP_{measured}$ is the total tea polyphenol content measured by wet chemical assay, $TP_{estimated}$ is the total tea polyphenol content derived using the inverted tea-PROSPECT model with the look-up table, and n is the number of observations in the sample.

Comparing the retrieval accuracies using empirical and physical approaches

Statistical approaches (partial least squares regression) have been successfully applied to predict tea polyphenols of fresh tea leaves of one tea cultivar from spectral data (Bian et al. 2010). To compare this empirical based model with the new tea-PROSPECT model, both models were used to estimate the concentration of total tea polyphenols. Two datasets of tea leaf samples, including young and old leaves of the same tea cultivar (as used for establishing the empirical model), as well as leaves from six different cultivars collected in the field, were examined. Then the relative RMSE and r^2 between measured and estimated tea polyphenols generated by the statistical and the tea-PROSPECT models were compared.

5.3 Results

Table 5-2 details the change in leaf biophysical and biochemical measurements for one tea cultivar over seven phenological stages, as well as for six tea cultivars. All the biochemical values are expressed in μg cm^{-2} or g cm^{-2}, being equivalent to the specific leaf area index. These values are within the range defined in Table 5-1. From young (foliar phenological stage *a*) to old (foliar phenological stage *g*) tea leaves, the tea polyphenol concentration decreased from 0.0023 g cm^{-2} to 0.0012 g cm^{-2}. Among the different tea cultivars, the Huang dan cultivar had the highest concentration of tea polyphenols (0.0029 g cm^{-2}), while the Fuding dabai cultivar had the lowest value (0.0015 g cm^{-2}).

Table 5-2 Leaf biophysical and biochemical measurements. All results are means of four duplicate field samples.

One tea cultivar, seven stages						
*Foliar Phenological stage**	*n*	*CHL** ($\mu g\ cm^{-2}$)	*Water* ($g\ cm^{-2}$)	*Tea polyphenols* ($g\ cm^{-2}$)	*Dry matter* ($g\ cm^{-2}$)	*SLA**** ($cm^2\ mg^{-1}$)
a	4	15.39	0.0222	0.0023	0.0083	0.1212
b	4	23.02	0.0176	0.0017	0.0075	0.1334
c	4	38.32	0.0180	0.0015	0.0073	0.1362
d	4	51.19	0.0194	0.0016	0.0087	0.1148
e	4	64.79	0.0194	0.0015	0.0094	0.1069
f	4	61.39	0.0163	0.0012	0.0081	0.1241
g	4	58.93	0.0161	0.0012	0.0082	0.1213
Six tea cultivars						
Cultivar name	*n*	*CHL** ($\mu g\ cm^{-2}$)	*Water* ($g\ cm^{-2}$)	*Tea polyphenols* ($g\ cm^{-2}$)	*Dry matter* ($g\ cm^{-2}$)	*SLA**** ($cm^2\ mg^{-1}$)
Fuding dabai	4	55.09	0.0135	0.0015	0.0086	0.1158
Tai cha	4	67.21	0.0095	0.0024	0.0116	0.0865
Mei zhan	4	96.45	0.0218	0.0022	0.0102	0.0980
Huang dan	4	69.64	0.0175	0.0029	0.0107	0.0931
E cha 1#	4	68.85	0.0158	0.0017	0.0093	0.1080
Fuyun 6#	4	81.02	0.0161	0.0023	0.0126	0.0796

**a-g* represent the seven different phenological stages from the top first leaf through to the seventh leaf; **CHL is total chlorophyll *a*+*b*; ***SLA is the specific leaf area.

Figure 5-1 illustrates the measured and simulated (through original PROSPECT and tea-PROSPECT models, respectively) reflectance for the seven phenological stages (*a*-*g*) of the Fuding dabai tea cultivar. To test the performance of the improved model for tea, the histogram presents the RMSE between the measured leaf reflectance and the values simulated using the tea-PROSPECT model, as well as using the original PROSPECT model without considering the contribution of total tea polyphenols. Compared to the original model, the tea-PROSPECT model yielded a higher accuracy with lower RMSE values. Except for the first phenological stage (*a*), the spectral RMSEs using the tea-PROSPECT model were below 0.05.

Figure 5-2 illustrates the measured and simulated leaf spectral reflectance for six tea cultivars. The improved tea-PROSPECT resulted in a more accurate spectrum reconstruction for each tea cultivar (with an average spectral RMSE of 0.025), than the original model did.

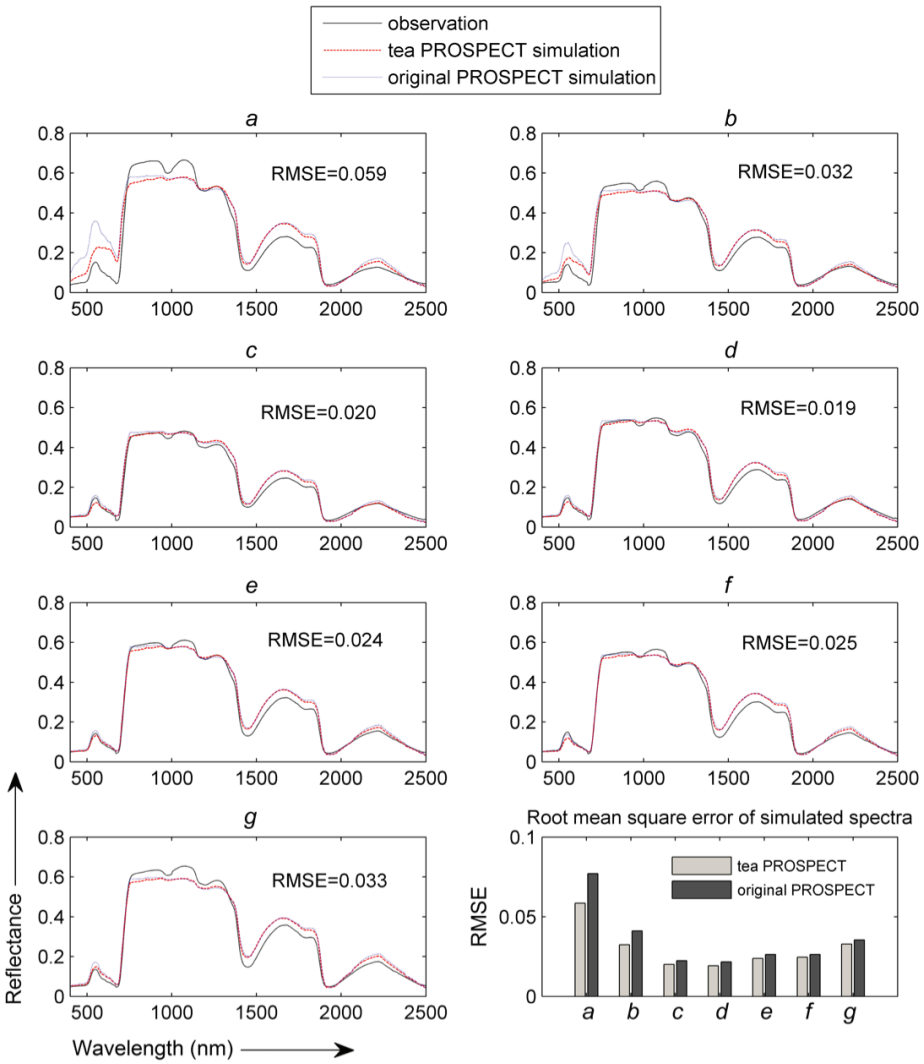

Figure 5-1 For the seven phenological stages (*a*-*g*) of one tea cultivar, the performance of the improved tea-PROSPECT model and the original PROSPECT model on the leaf spectral simulation. The spectral RMSE between the measured and the simulated spectra by the tea-PROSPECT model is shown on graphs *a* - *g*.

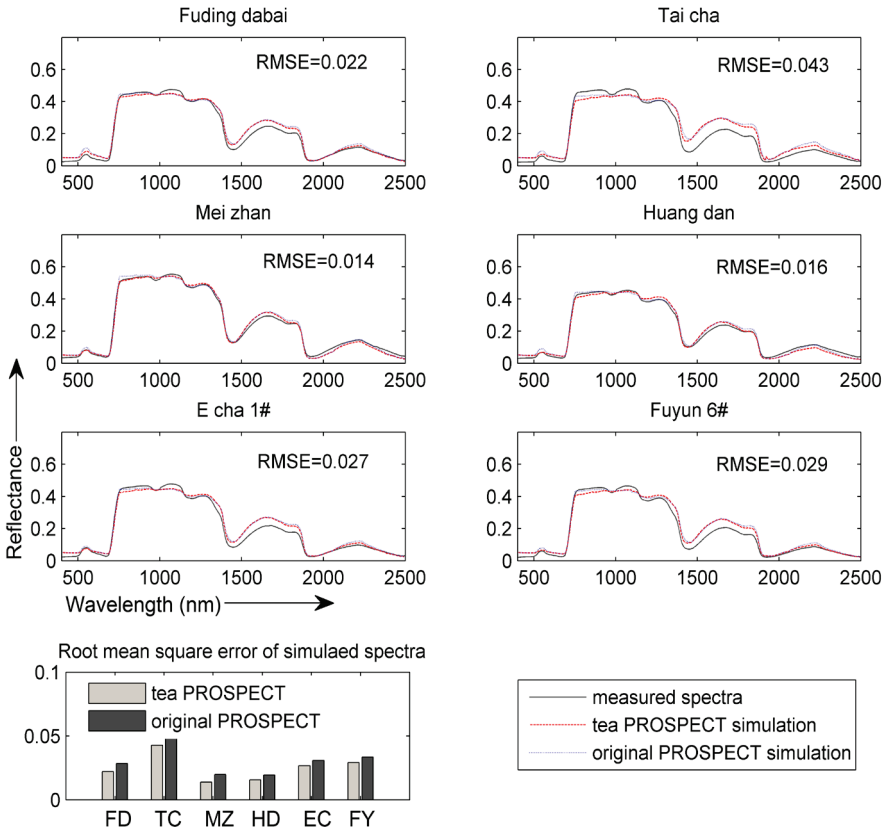

Figure 5-2 The performance of the improved tea-PROSPECT model and the original PROSPECT model on leaf spectral simulation for the six tea cultivars abbreviated as FD, TC, MA, HD, EC, and FY, respectively. The spectral RMSE between the measured and the simulated spectra by the tea-PROSPECT model is displayed on each graph.

Table 5-3 shows the retrieval accuracies after inversion of the tea-PROSPECT physical model, using fresh tea leaves of both varying ages and different tea cultivars. Based on the criterion of lowest RMSE of the measured spectra versus the spectra generated through the look-up table, the optimal combination of model parameters was chosen. For the physical approach, the lowest retrieved accuracy of total tea polyphenols was achieved from tea leaves at different phenological stages ($r^2 = 0.52$, RMSE/mean = 0.23), while the highest accuracy was achieved for varied tea cultivars ($r^2 = 0.66$, RMSE/mean = 0.10). For the two datasets together, the retrieval of tea polyphenols resulted in an r^2 equal to 0.69 with a relative RMSE of 0.17, which varied little from the results for the

different tea cultivars. The scatter plots of the relation between measured and estimated concentrations of tea polyphenols can be seen in Figure 5-3.

For comparison, table 3 lists the prediction results for tea polyphenols using the reported empirical regression model (Bian et al. 2010). As the empirical model has been established using the same Fuding dabai tea cultivar, the prediction accuracy for the tea samples of the seven different phenological stages was still high, with an r^2 of 0.86 and a relative RMSE less than 10%. However, when we extended the statistical correlation to the other five tea cultivars, the prediction accuracy decreased sharply to an r^2 of 0.19.

Table 5-3 r^2, RMSE, and relative RMSE between measured and estimated biochemical concentration of tea based on empirical and physical models.

	No. of observations	*r^2*	*RMSE (g cm $^{-2}$ x 10^{-3})*	*RMSE /mean*
Empirical approach (Bian et al. 2010)				
One tea cultivar (A)	7	0.86	0.13	0.08
Six tea cultivars (B)	6	0.19	0.51	0.23
Tea-PROSPECT model				
One tea cultivar (A)	7	0.52	0.37	0.23
Six tea cultivars (B)	6	0.66	0.23	0.10
Combined A and B	13	**0.69**	**0.31**	**0.17**

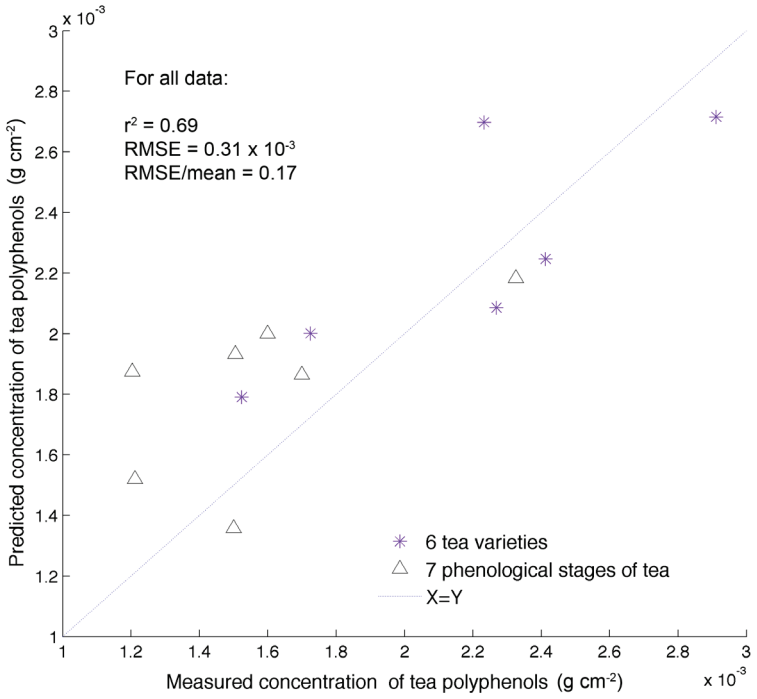

Figure 5-3 Predicted versus measured concentrations of total tea polyphenols, using the criterion of minimum spectral RMSE.

5.4 Discussion

5.4.1 Leaf spectral simulation of tea

The high concentration of tea polyphenols (18-36% of dry mass), and the importance of polyphenols for the quality of tea, stimulated us to explore reflectance spectroscopy as a method to predict leaf polyphenols. An improved leaf radiative transfer model incorporating polyphenols (tea-PROPSECT), and more accurately simulating the spectral reflectance of fresh tea leaves, is presented. The root mean square error between the simulated and measured spectral reflectance of tea leaves is generally less than 0.05 both for different phenological stages (Figure 5-1), as well as for different tea cultivars (Figure 5-2), indicating a good spectrum reconstruction by the tea-PROSPECT model (Jacquemoud et al. 1996). The spectral simulation results show that tea polyphenols indeed influence the absorption characteristics of tea leaves.

In contrast to the original PROSPECT model, in which only total chlorophyll, water and dry matter are included, the accuracy of the improved tea-PROSPECT model was increased by simulating the leaf optical properties of tea (Figure 5-1 and Figure 5-2). This may be attributed to the newly incorporated parameter of total tea polyphenols. The previous version of PROSPECT assumed that terrestrial plants all have similar foliar bio-chemistry (Feret et al. 2008), causing the derived parameters in PROSPECT to generalize for vegetation but lack specificity for certain plant targets. This finding is in agreement with the results of a previous study by Bian et al. (Bian et al. 2010), that the variation in tea polyphenol concentration is significantly related to the fresh tea leaf spectra.

The simulated reflectance corresponded with the measured reflectance for tea leaves at different growth stages (for the second through to the seventh leaf on one tea branch, RMSE<0.05). Of the seven phenological stages of tea, simulation of the first young tea leaves resulted in the highest error. As many single (top) first tea leaves were too small, we placed five or six leaves together for measuring leaf reflectance using the integrating sphere. The overlap of these leaves increased the leaf spectral reflectance in the near infrared plateau where the absorption features are largely influenced by the leaf mesophyll structure and vegetation biomass (Cho et al. 2007). As the PROSPECT model is designed to simulate the hemispherical reflectance and transmittance of a single leaf, the leaf structure parameter N for the young tea leaves at the top of the branch was overestimated. As a consequence, the predicted spectra did not closely match the measured spectra in the visible waveband region (Figure 5-1(*a*)).

From the histogram shown in Figure 5-2, it is clear that the spectral root mean square error between the measured and simulated spectra by the tea-PROSPECT model showed little variation from one tea cultivar to another. This result confirms the generality of the physically based leaf model.

5.4.2 Model inversion for the retrieval of total tea polyphenols

Although the tea-PROSPECT model performed well when synthesizing the whole leaf spectrum (400-2500 nm) for single and mixed tea cultivars, the retrieval of biochemical parameters by inversion of the model needs to be

further investigated (Jacquemoud et al. 1996). In this study, we focused on the estimation of total tea polyphenols, and for the first time, retrieved the concentration of total polyphenols. The retrieval of total tea polyphenols is feasible but resulted in relatively low accuracies (Table 5-3): for the two datasets of fresh tea leaves, one with different phenological stages, and one with different tea cultivars, the average r^2 equaled 0.59 and the RMSE values ranged from 10% (different tea cultivars) to 23% (different growing stages of one tea cultivar) compared to their average measured concentrations. This result confirms similar difficulties revealed in other studies in estimating biochemical parameters of vegetation (Darvishzadeh et al. 2008b), and may be explained by the influence of water in fresh leaves. Generally water represents around 70% of the fresh weight of a tea leaf. In the near infrared domain (beyond 1000 nm), water has strong absorption features which may mask minor absorptions caused by other biochemical components (Curran et al. 2001; Kokaly et al. 2009).

For the six tea cultivars, the retrieval accuracy of tea polyphenols (r^2 = 0.66, relative RMSE = 0.10) was relatively high compared to that for the fresh tea leaves at different phenological stages (r^2 = 0.52, relative RMSE = 0.23). A possible reason is that the lignin or cellulose content in tea leaves varies with leaf age, causing the spectral responses to the different concentrations of these biochemicals to overshadow the spectral signals that are related to the concentration change in total tea polyphenols, given that tea polyphenols and lignin have similar absorption features in the near-infrared region (Curran 1989; Soukupova et al. 2002).

Results in some studies demonstrated that the vegetation parameter could be estimated through inversion of physically based models (for example, PROSAIL) with accuracies comparable to those of statistical approaches (Darvishzadeh et al. 2011). However, for the biochemical parameter (total tea polyphenols) of tea, this is not the case. For one tea cultivar, the empirical approach performed much better (r^2 = 0.86, relative RMSE = 0.08) than the physically based model. However, when comparing the tea-PROSPECT model with empirical approaches, we have to make a tradeoff between accuracy and generality. When applying the empirical model for Fuding dabai tea with the mixed tea cultivars, the r^2 dropped sharply to 0.19 and the relative RMSE increased from 0.08 to 0.23. One of the possible reasons is that other tea

cultivars have a greater range of tea polyphenols compared with Fuding dabai (Table 5-2), resulting in the tea cultivars with very high or low concentrations not being well predicted.

When inverting the tea-PROSPECT model, the prediction accuracy was not as high as for the empirical approach. Nevertheless, it produced relatively stable results for the two datasets (different growing stages and different cultivars), and for the pooled datasets (r^2 = 0.69, relative RMSE = 0.17, Figure 5-3), which indicates that the physical model based inversion may be more accurate when generalized and applied to another area (Cho 2007).

5.5 Conclusions

Based on the widely used PROSPECT model, we proposed an improved leaf radiative transfer model for tea by introducing tea polyphenol to the list of parameters,. In terms of reflectance reconstruction, a low spectral root mean square error was obtained not only for one tea cultivar, but also for different tea cultivars, which demonstrates the capability of the improved tea-PROSPECT model to accurately simulate the whole leaf spectrum (400-2500 nm) of the tea species. As the tea polyphenol total has been included in the model, its concentration is thus retrievable. Although the biochemical retrieval resulted in a lower prediction accuracy, compared to the empirical approach the physical model appeared more robust. This result indicates the potential of model inversion (through a look-up table) for retrieving tea polyphenols using hyperspectral measurements.

Overall, the improved tea-PROSPECT model aids our understanding of the effects of foliar biochemicals on tea leaf optical properties. As prior knowledge is important when refining model parameters, and increasing the retrieval accuracy, this model will be further validated and refined for tea plants using widely different cultivars and growing conditions, to finally be combined with a canopy reflectance model to estimate the quality of tea using hyperspectral remote sensing in space and time.

CHAPTER 6

SYNTHESIS:

ASSESSING THE QUALITY OF TEA FROM LABORATORY NIRS TO FIELD SPECTRAL MEASUREMENTS

6.1 Introduction

Higher crop yields are essential, but it is not enough just to grow more. Higher quality is important to maximize the value of the crop and to meet rising consumer expectations (Gebbers and Adamchuk 2010; Singhal et al. 1997). For that reason, increasing attention has been drawn from yield estimation to quality estimation of crops and forage plants using hyperspectral remote sensing data (Kokaly et al. 2009; Skidmore et al. 2010; Zarco-Tejada et al. 2004). Tea (*Camellia sinensis* (L.)) is an important economic crop and the market price of tea depends largely on its quality (Dutta 2011; Yan 2007). Therefore, if it is possible to monitor the quality of tea in large scale using remote sensing techniques, one can estimate the economic benefit of the target tea plantation before tea harvesting. In this study, for the first time, the concentrations of biochemical compounds as indicators of the quality of tea were estimated using reflectance spectroscopy systematically from dried tea powders, via fresh tea leaves to tea plants' canopies, trying to explore the potential of hyperspectral remote sensing on tea quality estimation.

Research in the tea industry has revealed that the total tea polyphenols, free amino acids, soluble sugars and caffeine, from among the hundreds of chemical compounds found in tea contribute the most to the taste of the tea and its quality (Scharbert et al. 2004; Sumpio et al. 2006; Yamamoto et al. 1997). For example, Figure 6-1 shows a simplified correlation between tea sensory preferences and the chemical components of green tea when considering the concentrations of the total tea polyphenol and free amino acids, which is drawn from a published dataset by Li (2002). Significant relationships have also been identified between the soluble sugars and caffeine and tea sensory scores (Dutta 2011; Gu et al. 2005; Nakagawa 1975). Thus, I focused on assessing these four biochemical parameters using hyperspectral techniques in this research.

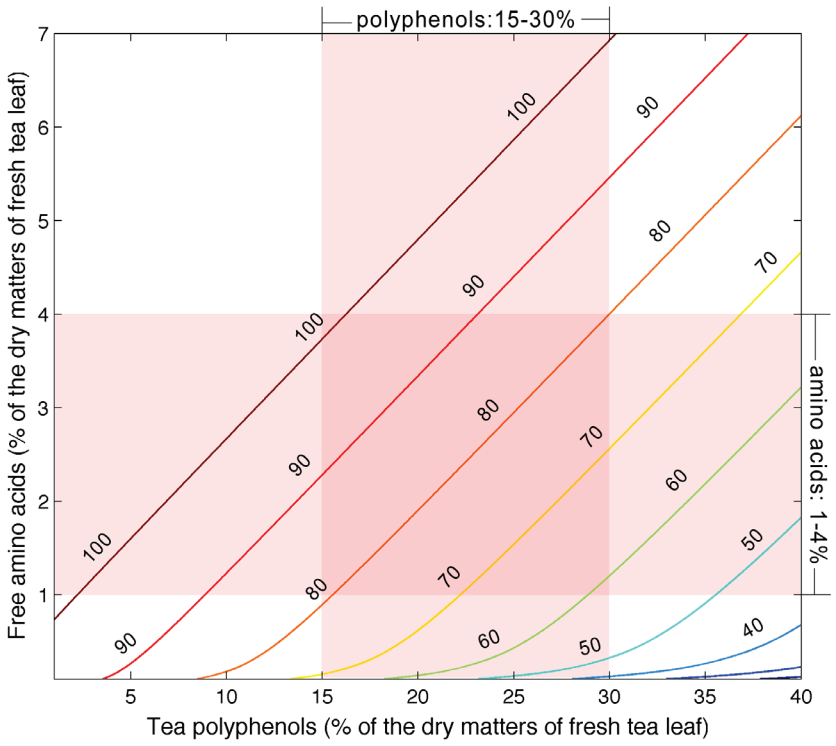

Figure 6-1 Sensory score of green tea as a function of total tea polyphenols and free amino acids, based on Li (2002). The number above each contour line represents the taste score with the corresponding chemical concentrations of tea. The two red belts show the normal concentration ranges of the two biochemicals for green tea. We should note that this interpolation is just a simplification of the reality: as some articles have pointed out, when the concentration of polyphenols is lower than a certain amount, the signature style of tea's bitter taste is absent, and so the sensory evaluation scores go down as the tea flavour is not there anymore.

In general, the current remote sensing approaches for estimating biochemical attributes of vegetation include statistical (inductive) and physically based (deductive) models (Atzberger 2004; Darvishzadeh et al. 2008b; Kokaly et al. 2009). Each of them has pros and cons, and has been applied to achieve different purpose (Asner 1998; Cho and Skidmore 2006; Darvishzadeh et al. 2008b; Ferwerda and Skidmore 2007; Mutanga and Kumar 2007). Chapters 2, 3 and 4 of this thesis presented empirical models that were built for estimating the foliar biochemical parameters associated with tea's quality at different stages (powder, fresh leaf, and tea plant) for one or mixed tea cultivars, using

statistical approaches. Chapter 5 presented an improved leaf optical model incorporating the total tea polyphenols, developed with the further aim of accurately estimating the tea quality for any tea cultivars and growing stage at the regional scale.

This final chapter summarizes the main results of this study and their interrelationships (Figure 6-2), in order to gain a better understanding of the similarity and difference of the results and methods used when retrieving biochemical compounds of interest from laboratory to field level. Also, the practical relevance of this thesis for managing tea plantations, and the contributions of this thesis to precision agriculture and food studies will be emphasized. Finally, directions for future research based on current progress will be addressed, including extending this study to larger scales and monitoring tea quality in the context of global changes.

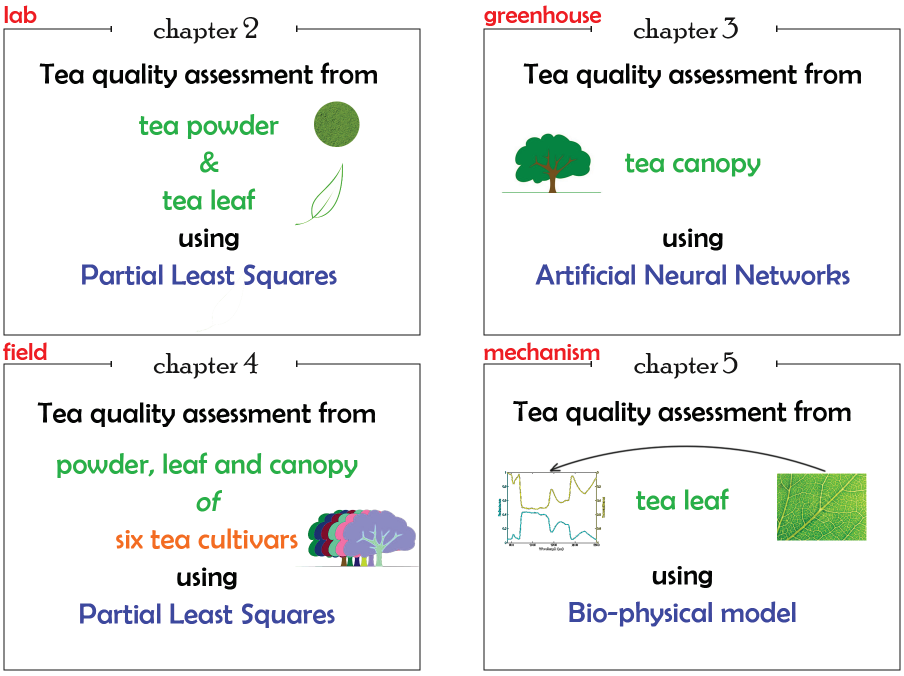

Figure 6-2 The context of and the relations between the different thesis chapters.

6.2 Laboratory level

- Can the chemical composition (quality-related biochemicals) of tea leaves be detected from the spectral reflectance of fresh tea leaves? (for one tea cultivar)?

Accurate remotely sensed estimates of the foliar biochemical concentration of tea plants' canopies can provide a valuable aid to the prediction of the quality of tea before plucking over a wide range of scales. The development of a hyperspectral remote sensing technique with a large number of narrow, contiguously spaced spectral bands has provided the possibility of retrieving biochemical parameters in vegetation (Knox et al. 2011; Lee et al. 1994; Yue et al. 2008). Based on hyperspectral data, a lot of biochemical contents including chlorophyll, nitrogen, water, lignin, starch, cellulose, protein, and some trace elements such as potassium, sodium, and calcium have been estimated with different accuracies (Curran et al. 1992; Kokaly and Clark 1999). Most of the research on the detection of biochemical content has focused on grass, forestry, and agricultural crops such as wheat (Darvishzadeh et al. 2008c; Haboudane et al. 2002; Hansen and Schjoerring 2003; Mutanga et al. 2004b). The use of this biochemical content for tea quality prediction has received little attention (Dutta et al. 2011; Ishikawa et al. 2006).

In order to use imaging spectrometry to assess tea quality, it is required to analyse the spectral characteristics of quality-related biochemical compounds of tea, and establish models to predict the content of these biochemicals. The complexity of the spectroscopy changes with the form of the plant material used —it typically increases as one passes from the ground powder of dried leaves, via fresh whole leaves, to canopies (Curran 1989; Kokaly et al. 2009; Majeke et al. 2008; Mutanga et al. 2009). Therefore, most of those studies that try to explore the spatial biochemical variation of vegetation start from the laboratory, with dried leaf powders or fresh leaf materials, in order to simplify the experiment and reduce the number of variables. After satisfactory results have been obtained at this level, these attempts can be extended to the canopy level (Curran et al. 2001).

Laboratory near infrared spectroscopy (NIRS) was first used in agricultural applications by Norris (1964) to measure the moisture in grain. Since then, this technique has been used for the rapid analysis of mainly moisture, protein and

the fat content of a wide variety of agricultural and food products (Nicolaï et al. 2007). For tea, reflectance spectroscopy has been used to detect alkaloids and phenols using laboratory NIRS since the 1990s (Schulz et al. 1999). Their results showed a high correlation between tea-quality related biochemicals and reflectance spectroscopy. But these studies were limited to the dried and ground powder of tea leaves (Bian et al. 2010; Chen et al. 2008a). Research on fresh leaves is of great importance, because it is a step towards the application of imaging spectroscopy to the prediction of tea quality. In Chapter 2, it was therefore hypothesized that the two main biochemical tea quality indicators, total tea polyphenols and free amino acids, can be detected from the spectral reflectance of fresh tea leaves. For comparison, experiments were also designed for dried tea leaf powder using laboratory NIRS.

In order to maximize the range of foliar biochemical concentration (Ferwerda et al. 2006a), tea leaf samples of the tea cultivar 'Fuding dabai' were collected at different phenological stages. In the experimental field, using a single season's growth, leaves were picked from a branch, starting from the first leaf with young buds and continuing to the seventh leaf. The collected leaf stacks had their spectra measured using an ASD spectrometer in a dark-room laboratory, then the leaves were dehydrated and ground for NIRS measurements. Compared to narrow band vegetation indices such as the normalized difference vegetation index (NDVI) and the red-edge position (REP), partial least squares regression (PLSR) can exploit more of the rich information content of hyperspectral data (Cho et al. 2007). Thus, in Chapter 2, partial least squares regression was performed to establish the relationship between the reflectance and the biochemical content for leaf powders as well as fresh leaves.

Compared to the results obtained at the dried powder level, the prediction accuracy decreased for both of the two biochemicals at the leaf level. However, satisfactory accuracy was still found for fresh tea leaves with a cross-validated r^2 of 0.91 for total tea polyphenols and 0.93 for free amino acids, and the root mean square errors of the cross validations were 13.74 mg g^{-1} and 2.32 mg g^{-1}, respectively. The results of this experiment suggest that the hyperspectral technique is an effective way to predict the concentrations of total tea polyphenols and free amino acids in fresh tea leaves, which are major indicators of tea quality. The important wavelength regions for the prediction of these two

biochemicals were selected and it was found that these regions changed from powder to fresh leaf level.

6.3 Field level

- How can the chemical composition of young tea leaves be inferred from the spectroscopy of the tea canopy? (for one tea cultivar)?

At the leaf level, the reflectance variability at a given wavelength is mainly a function of leaf structure, concentration of biochemicals, and water content (Curran et al. 1992; Fourty et al. 1996; Gates et al. 1965). Compared to the fresh leaf level, the canopy level is even more complicated, since other factors, such as the soil background and canopy structure, influence the measured radiance signal (Dawson 2000; Elvidge 1990; Huete 1989; Peterson et al. 1988). Although satisfactory results of prediction have been achieved when using the reflected spectra of fresh tea leaves in the laboratory (Chapter 2), the possibility of using hyperspectral data for the estimation of the foliar biochemicals of tea at the canopy level in the field needs to be further proved (Chapter 3 and 4).

The study showed that partial least squares regression is an effective method for retrieving the biochemical parameters from the spectral reflectance of both fresh leaves and leaf powder. However, a major drawback of this method is that the relationship between the independent (wavebands) and dependent (biochemicals) variables is assumed to be linear (Cho et al. 2007; Hansen et al. 2002), though this may not be the case and needs to be explored. In this regard, there is interest in establishing a biochemical–spectral relationship using non-linear regression methods, such as the artificial neural networks (ANN) model (Atkinson and Tatnall 1997; Mutanga and Skidmore 2004b; Si et al. 2012).

Using hyperspectral data from the tea canopy level, a new hybrid approach involving artificial neural networks and a successive projections algorithm was developed in Chapter 3 for estimating tea foliar biochemical concentrations, namely total tea polyphenols, free amino acids, soluble sugars, and caffeine. The quality of fresh tea leaves is closely related to the niche conditions where the plant grows, including irrigation, sunlight, and temperature. Especially, the nitrogen, phosphorus, and potassium content of the soil nutrients has a direct impact on the foliar chemical composition of tea, thus affecting the tea's taste and quality (Chong et al. 2009; Xu et al. 2008). To control the growing

conditions of the tea plants, a greenhouse experiment was conducted: young tea plants of the same cultivar as studied in Chapter 2 were grown in a greenhouse with varied nitrogen, phosphorous, and potassium content in the soil. After 4 months of growth, canopy spectral measurements were taken on each pot of tea plants in sunny and cloud-free day in field conditions, to investigate whether the tea quality as a response to different soil nutrient conditions can be detected using *in-situ* canopy hyperspectral data.

In this chapter, the successive projections algorithm was applied to select the optimal wavelength bands for the input data of a neural network model for the prediction of the four biochemicals of interest. For comparative purposes, other commonly used band selection methods including principal component analysis and spectra random selection were also applied to reduce the hyperspectral data before running the neural network. The results showed that the novel integrated approach proposed here led to the highest prediction accuracy for each of the four biochemicals: based on an independent test dataset, the best-trained neural networks resulted in a coefficient of determination (r^2) between predicted and observed concentrations of 0.82, 0.76, 0.76 and 0.74 for tea polyphenols, free amino acids, soluble sugars, and caffeine, respectively; for the first three biochemicals, the root mean square error of prediction (RMSEP) is less than 5 percent of the mean value.

It can be concluded that based on the optimal number of wavelengths selected by the successive projections algorithm, the neural network trying to simulate possible complex nonlinearity in the data can be applied to a better estimation of the concentration of tea polyphenols, free amino acids, soluble sugars, and caffeine. However, we should note that, this experiment was undertaken in a greenhouse with intentionally contrasted soil conditions, and only one tea cultivar was involved. The question of whether the quality parameters can be estimated from the tea canopy spectra for different tea cultivars under natural condition, was explored in Chapter 4.

6.4 From laboratory to field

- *(1) How does the retrieval capability as one passes from dried tea leaf powder, via fresh whole leaves, to canopies, and (2) Are the spectral bands*

essential for chemical retrieval at one scale still important at another scale (across mixed tea cultivars)?

Laboratory spectroscopy has been developed, tested, and used to analyse the plant biochemicals of both dried and ground samples (Yoder and Pettigrew-Crosby 1995). The use of spectroscopy to predict the chemical composition of living plants is still in an experimental phase, for the leaf water, canopy structural variables, and other environmental factors can influence the spectral characteristics (Curran et al. 2001; van der Meer 2004). Although several studies have recently shown the potential of hyperspectral techniques for detecting the chemical composition of vegetation at the canopy level in the field (He et al. 2011; Martin and Aber 1997; Mutanga and Skidmore 2004a; Treitz et al. 2010), only a few studies have compared the results from dried leaf powder with those from fresh leaves and from canopies.

Chapters 2 and 3 provided the necessary methods for deriving biochemical information using tea spectra at each scale, which contributed to the estimation of tea quality for one tea cultivar. Chapter 4 focused on building an empirical correlation between the tea-quality related biochemicals and the spectral reflectance across different tea cultivars for tea powders, fresh leaves, and plants in the tea garden (Figure 6-3). This chapter aims at exploring the potential of hyperspectral remote sensing for assessing tea quality in terms of tea species, and obtain a general idea of how the retrieval capability and influential wavebands used for the prediction change from indoor laboratory to the field. The canopy spectra of six different tea cultivars, namely Fuding dabai (FD), Fuyun 6# (FY), E cha 1# (EC), Tai cha 12# (TC), Huang dan (HD), and Mei zhan (MZ), in the tea garden were measured using ASD field spectrometer. Consequently, the top young tea leaves were collected and sent to the laboratory for spectral measurement on fresh leaf and dried leaf powders, using an ASD spectrometer. Although the results in Chapter 2 and 3 suggested that both a partial least squares regression (a linear modeling approach) and an artificial neural network (a non-linear modeling approach) are effective methods for retrieving the foliar biochemical parameters from the spectral reflectance, partial least squares regression has been adopted as the regression method in this chapter because it is a deterministic algorithm, thus the results are more comparable.

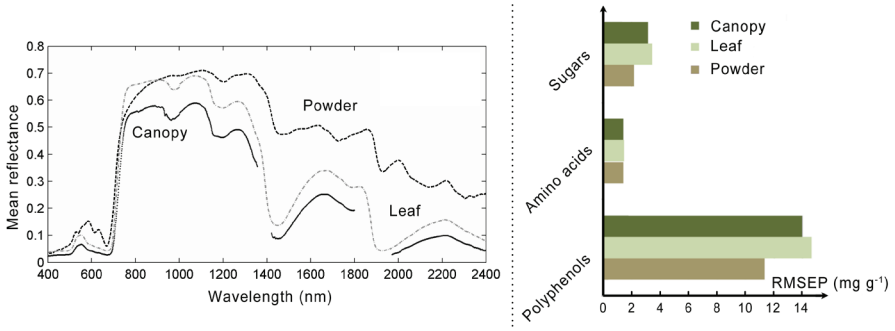

Figure 6-3 Derived from Chapter 4, this shows the measured reflected spectra (mean values of 48 samples), and the prediction accuracy (RMSEP, root mean square error of prediction) for tea powders, fresh tea leaves and canopies of tea plants.

For the first time, we demonstrated the utility of hyperspectral techniques for predicting tea-quality related biochemicals at three different levels. As we expected, the powder level produced the highest accuracy for the estimation of each biochemical component, while an important finding is that the prediction results at the canopy level in the field slightly outperformed the results from the leaf level (the leaf spectra were measured in a dark room with controlled conditions). This finding is supported by Asner and Martin (2008) who found that at the canopy level, partial least squares regressions at constant leaf area index levels and homogenous canopy structure are equally or more robust than those made at the leaf level. This result gives rise to positive expectations from airborne and spaceborne sensors for repeated mapping of tea quality, at the landscape or the regional scale. The results also proved that the spectra–chemical relationship exists not only for a single tea cultivar but also for an entire tea species. The first question raised at the beginning of this section has now been answered.

After analysing the important wavebands that the partial least squares algorithm used to predict the biochemical compounds, I found that the wavelengths that were considered the most important were mostly different at the three levels. These wavelengths have been classified into two categories: ‘direct’ for those bands that are located near the absorption peaks of the biochemicals of interest, and ‘indirect’ for those bands located in other places of the spectrum. It was found that at the lower levels (powder and leaf levels), those ‘direct’ bands contribute more to the prediction, while at the higher level, the canopy level, the

regression made use of more 'indirect' bands to predict the biochemical concentrations. This finding helped us to answer the second question.

6.5 Leaf optical model for tea species

– (1) For further estimation of tea quality on a larger scale, can a physically based model be developed, incorporating the foliar biochemicals as quality indicators of tea? (2) What is the retrieval accuracy when the model is inverted?

To date, the focus of many studies has been on empirical methods for estimating biochemical or structural variables of vegetation (Adam et al. 2010; Cho et al. 2007; Darvishzadeh et al. 2011; Skidmore et al. 2010). Although empirical models are relatively simple to build, the accuracy of the models is hard to maintain across different samples. On the other hand, a physically based model is able to represent the relationship with generality; it has an understanding of the physical process behind the relationship. Thus, Cho et al. (2007) have pointed out that the future of hyperspectral remote sensing might hinge on enhancing the link between empirically and physically based approaches.

When using statistical approaches, the previous three chapters 2, 3 and 4 showed that the concentration variation of tea-quality related biochemicals was significantly related to the fresh tea leaf spectra or the canopy spectra, for either one tea cultivar or mixed cultivars. For further estimation of tea quality on a larger scale, Chapter 5 tried to develop a physically based model incorporating the foliar biochemicals as quality indicators of tea.

The optical properties of a plant canopy largely depend on the optical properties of leaves and soil background (Botha et al. 2010). An improved leaf optical model for tea was developed, based on an updated version of the PROSPECT model which has been validated and applied in numerous studies (Feret et al. 2008; Fourty et al. 1996; Jacquemoud et al. 2009). For tea leaves, total tea polyphenols make up more than 20 percent of its dry mass, while in most other plants, polyphenols that belong to secondary metabolites exist at low concentrations. Therefore, the tea polyphenols have to be taken into account when modeling the reflectance/ transmittance spectrum of tea leaves to ensure an accurate prediction. The improved tea-PROSPECT model takes tea polyphenols separately as a specific input parameter, as shown in Figure 6-4.

Compared to the original leaf optical model, the presented tea-PROSPECT model more accurately simulates the spectral reflectance of fresh tea leaves, both for different phenological stages of the tea leaves and for different tea cultivars, indicating a good spectrum reconstruction by this model. To validate the utility of the model, the retrieval of the biochemical parameters (e.g., the content of total tea polyphenols) by inversion of the model was also investigated, and compared with the empirical model established before for one tea cultivar. The results show that although the empirical approach performed better than the physically based model for one tea cultivar, the prediction r^2 dropped quickly when it was applied to mixed tea cultivars. The improved tea-PROSPECT model retrieves, with a moderate accuracy, the concentration of tea polyphenols from the spectra for any growth stage or any cultivar, thereby demonstrating its generality and robustness across a wide range of conditions.

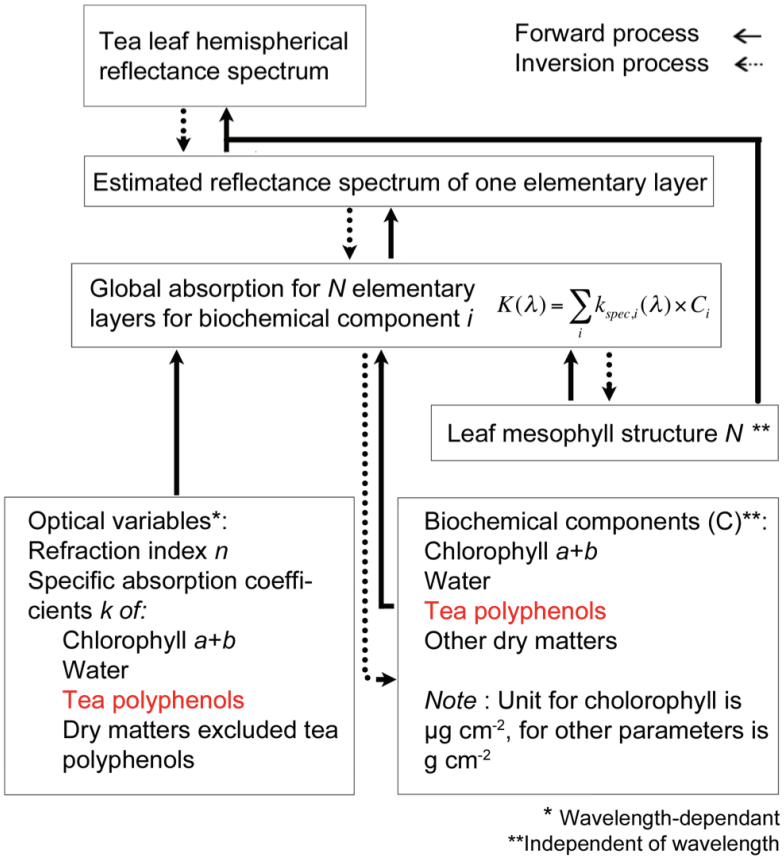

Figure 6-4 Flowchart of the steps followed by the improved tea-PROSPECT model incorporating total tea polyphenols' content as a new input parameter. This was redrawn based on the work of (Botha et al. 2006).

6.6 Practical relevance of this research

It is crucial to evaluate not only the quantity but also the quality of economic crops in space and time, for the quality is an essential factor influencing their products' quality and the market value (Maeda et al. 2009; Williams et al. 1988; Yan 2007). Tea is a global cash crop, and its quality is generally appraised by professional tea tasters after plucking and processing (Liang et al. 2008). My study proposed an objective and non-destructive method of hyperspectral remote sensing to obtain qualitative information about tea during its growing period. This method based on reflectance spectroscopy at the powder, leaf, and canopy levels, can be transferred to imaging spectroscopy to predict the quality

of tea in large areas, and help monitor tea quality status, which is important information for tea farmers and the tea industry.

6.6.1 For tea plantations managers (individual plantations)

The results from chapters 2 and 3 show the potential of hyperspectral remote sensing for the estimation of the biochemical parameters of tea at the stage of fresh leaves. This can help tea garden managers to predict the quality of tea before plucking and evaluate its market value in advance.

Especially, the results obtained from Chapter 3 implies that the biochemical concentrations, which vary with different soil nutrient levels, could be detected through tea canopy spectral reflectance. This may help managers respond more quickly and adjust fertilizer levels (Goel et al. 2003) to improve tea quality during the growing season.

6.6.2 For managers of tea plantations and processing plants

Most studies that have been carried out previously predict the productivity of tea or discriminate tea cultivars at the landscape or regional scale (Li and He 2008; Rajapakse 2002). In this research, for mixed tea cultivars, the plant's biochemical composition could be statistically correlated with the hyperspectral data (Chapter 4). An improved tea-PROSPECT model, which is more robust and portable, has been developed and could be combined with the canopy model for retrieving the biochemical parameters as indicators of tea quality (Chapter 5). Based on the information about tea quantity and quality at the local or regional scale, we can accurately estimate the economic value of the tea in one area. Also, based on the spatial pattern of the tea quality at the regional scale, a local administrator can make spatially explicit suggestions on tea cultivation, to boost the regional profit of the tea industry.

6.6.3 For Ministry of Agriculture (Domestic tea production and international tea trade)

With the quality information of tea plantations retrieved from remote sensing data, the government can classify and rank tea production areas within a country, and make more reasonable and practical strategies for tea market segmentation (Dutta 2011; Moran et al. 1997; Pinter et al. 2003). Furthermore, with the capability of obtaining both the productivity and quality information on global

tea agriculture, the government can make better predictions of global tea prices and enhance the competitiveness of its native production, better avoiding risk.

6.6.4 Contributions to agricultural and food studies

Precision farming or precision agriculture is a farming management concept that requires obtaining quantitative information about the agriculture in specific area using the technology of the global positioning system (GPS), remote sensing and the geographic information system (GIS) (Gebbers and Adamchuk 2010). However, previous studies of remote sensing have mainly focused on retrieving biophysical parameters (such as LAI and biomass), or chlorophyll content, which can only reflect the health status of the plants (Blackburn 2007; Cho et al. 2007). The present study aimed at retrieving the biochemical concentrations including tea polyphenols, free amino acids, and so on, and the results showed that quantitative information on the quality and taste of tea can be achieved quickly and non-destructively through hyperspectral techniques. This will be useful to further develop aspects of precision agriculture for tea plantations.

This study can also be seen as an extension of food study: for the first time, the quality of tea is assessed outside the laboratory, where wet chemistry and NIRS techniques have been applied, through estimating the biochemical components as quality indicators of tea in fresh tea leaves (the raw material of tea products) using hyperspectral techniques.

6.7 General conclusion

The main aim of this thesis was to investigate whether the biochemical concentrations of tea-quality indicators, namely total tea polyphenols, free amino acids, soluble sugars, and caffeine, can be estimated quantitatively and non-destructively using hyperspectral techniques. Passing from the laboratory NIRS level to the field canopy level, I have examined the performance of different statistical techniques, including linear (partial least squares regression) and non-linear modeling approaches (a hybrid approach of artificial neural networks with a successive projections algorithm) for assessing the quality of tea. Furthermore, the study proposed a physically based model—the tea-PROSPECT model—to accurately simulate tea leaf spectra. This improved model incorporates the total tea polyphenols as a new parameter, and the

performance of this physical model for retrieving the foliar biochemicals of tea were investigated and compared with the statistical approach. The general conclusion is that the four biochemicals of interest can be significantly correlated with the hyperspectral data from powder, leaf, and canopy. In particular, for total tea polyphenols, something which has high concentrations in the tea leaf (18%–36% of the dry mass), the improved tea-PROSPECT model could accurately retrieve its concentration from the leaf spectra for varied growth stages and varied cultivars. If we achieve the prediction of tea quality using remote sensing images, the technique developed in this thesis can be applied to hyperspectral remote sensing to predict the quality of tea in space and time.

6.8 Future perspectives

To explore the potential of hyperspectral remote sensing for assessing the quality of tea, several interrelated experiments were carried out in conditions ranging from the laboratory to the natural field conditions. However, this study did not cover more tea cultivars or take into account other quality-related biochemicals of tea. More importantly, the success at the canopy level has not been extended to the image level: the use of imaging spectroscopy for predicting the foliar chemistry of tea still faces many challenges. Also, the physically based model needs to be up-scaled to the air or space-borne level. These important issues need to be further investigated in order to finally achieve the ultimate objective of this study. The details of the avenues for future work relevant to this study are as follows:

1) Either biophysical or statistical modelling requires enough field survey to obtain sufficient prior knowledge or the collection of sufficient field data for building a more accurate model (Bouman 1992; Darvishzadeh et al. 2008b; Feret et al. 2008). Thus, the models built in this thesis will be further validated and refined for tea plants in widely different cultivars or growing conditions. Meanwhile, besides the main four biochemicals illustrated in this thesis, other biochemical components such as mineral elements and volatile aromatic chemicals will be included. These biochemicals are important for the tea infusion's taste and smell, thereby influencing the tea's quality (Gu et al. 2005; Yamamoto et al. 1997).

2) The final aim of the study is to assess the quality of tea using hyperspectral remote sensing in space and time. Thus, the future of the study lies in further extending the methods (already used and refined at the canopy level) to the regional or continental scale, for estimating the biochemical characteristics related to tea quality by air/space borne sensors. When using imaging spectroscopy, more factors, such as the soil background properties, the observation geometry, the atmospheric condition, and the shade trees in tea plantations, will influence the radiance signal received (Clark et al. 2011; Peng et al. 2012). The methods developed in this study may be employed after these disturbances have been taken into account. Regarding the physically-based leaf optical model for tea, on the one hand, the input parameters need to be investigated and adjusted based on prior knowledge obtained by field surveys and a literature review as well; on the other hand, the leaf model should finally be combined with a canopy reflectance model and an atmospheric radiative transfer model to estimate the quality of tea over a large area (Verhoef and Bach 2007).

3) Many environmental factors involving the topography, rainfall, temperature, frost-free period, etc., have a great impact on the biochemical composition of tea, thereby influencing the tea's quality (Owuor et al. 2011). In further study, the questions of "Why does regional/district/local quality vary for tea?" and "Can quality variation be mapped or monitored using remote sensing combined with other environmental variables such as climates, soils and so on?" need to be answered. By combining hyperspectral techniques and quantitative Geo-information, the suitable areas for different tea cultivars will be mapped and predicted.

4) Climate change is predicted to alter carbon dioxide levels, temperature, and precipitation patterns in the world (Dinar and Mendelsohn 2012). Under this circumstance, in the long run, the taste of the tea may change. To investigate this issue, we should firstly focus on the potential scenarios of the environmental changes that matter to tea plantations. For instance, a rise in temperature or a fall in the annual amount of solar radiation. Then, based on biological and botanical knowledge, the response of the tea plant in terms of quality can be modeled in a spatially explicit way. This prediction can be validated using the remote sensing method developed in this study at the global scale.

BIBLIOGRAPHY

Acland, J.D., 1971. *East African Crops,* Longman for the Food and Agriculture Organization of the United Nations, London, 252 p.

Adam, E., O. Mutanga, and D. Rugege, 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management,* 18(3): 281-296.

Agarwal, B., U. Singh, and M. Banerjee, 1992. In vitro clonal propagation of tea (*Camellia sinensis (L.) O. Kuntze*). *Plant cell, tissue and organ culture,* 30(1): 1-5.

Ainsworth, E.A., and K.M. Gillespie, 2007. Estimation of total phenolic content and other oxidation substrates in plant tissues using Folin-Ciocalteu reagent. *Nature Protocols,* 2(4): 875-877.

Araújo, M.C.U., T.C.B. Saldanha, R.K.H. Galvão, T. Yoneyama, H.C. Chame, and V. Visani, 2001. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems,* 57(2): 65-73.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment,* 64(3): 234-253.

Asner, G.P., J.A. Hicke, and D.B. Lobell, 2003. *Per-pixel analysis of forest structure: Vegetation indices, spectral mixture analysis and canopy reflectance modeling, Remote Sensing of Forest Environments: Concepts and Case Studies* (M. Wulder and S. Franklin, editors), Kluwer Academic Publishers, Boston, pp. 209-254.

Asner, G.P., and R.E. Martin, 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment,* 112(10): 3958-3970.

Asner, G.P., C.A. Wessman, C.A. Bateson, and J.L. Privette, 2000. Impact of Tissue, Canopy, and Landscape Factors on the Hyperspectral Reflectance Variability of Arid Ecosystems. *Remote Sensing of Environment,* 74(1): 69-84.

Astill, C., M.R. Birch, C. Dacombe, P.G. Humphrey, and P.T. Martin, 2001. Factors Affecting the Caffeine and Polyphenol Contents of Black and Green Tea Infusions. *Journal of Agricultural and Food Chemistry,* 49(11): 5340-5347.

Atkinson, P., and A. Tatnall, 1997. Introduction neural networks in remote sensing. *International Journal of Remote Sensing,* 18(4): 699-709.

Atoui, A.K., A. Mansouri, G. Boskou, and P. Kefalas, 2005. Tea and herbal infusions: their antioxidant activity and phenolic profile. *Food Chemistry,* 89(1): 27-36.

Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sensing of Environment,* 93(1): 53-67.

Atzberger, C., T. Jarmer, M. Schlerf, B. Kötz, and W. Werner, 2003, Spectroradiometric determination of wheat bio-physical variables: comparison of different empirical-statistical approaches. *Proceedings of the 23rd EARSeL Symposium: Remote Sensing in Transition,* 2-5 June, Ghent, Belgium, pp. 463-470.

Bajcsy, P., and P. Groves, 2004. Methodology for hyperspectral band selection. *Photogrammetric engineering and remote sensing,* 70(7): 793-802.

Balabin, R.M., R.Z. Safieva, and E.I. Lomakina, 2008. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. *Chemometrics and Intelligent Laboratory Systems,* 93(1): 58-62.

Banerjee, B., 1988. Dry matter production and partitioning by tea varieties under differential pruning. *Applied agricultural research,* 3(4): 326-328.

Baret, F., and T. Fourty, 1997. Estimation of leaf water content and specific leaf weight from reflectance and transmittance measurements. *Agronomie,* 17(9-10): 455-464.

Beeri, O., R. Phillips, J. Hendrickson, A.B. Frank, and S. Kronberg, 2007. Estimating forage quantity and quality using aerial hyperspectral imagery for northern mixed-grass prairie. *Remote Sensing of Environment,* 110(2): 216-225.

Bian, M., A.K. Skidmore, M. Schlerf, T. Fei, Y. Liu, and T. Wang, 2010. Reflectance spectroscopy of biochemical components as indicators of tea (*Camellia sinensis*) quality. *Photogrammetric Engineering & Remote Sensing,* 76(12): 1385-1392.

Blackburn, G.A., 1998. Quantifying chlorophylls and carotenoids at leaf and canopy scales - an evaluation of some hyperspectral approaches. *Remote Sensing of Environment,* 66(3): 273-285.

Blackburn, G.A., 2007. Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany,* 58(4): 855-867.

Borse, B., L. Jagan Mohan Rao, S. Nagalakshmi, and N. Krishnamurthy, 2002. Fingerprint of black teas from India: identification of the regio-specific characteristics. *Food Chemistry,* 79(4): 419-424.

Botha, E., B. Leblon, B. Zebarth, and J. Watmough, 2010. Non-destructive estimation of wheat leaf chlorophyll content from hyperspectral measurements through analytical model inversion. *International Journal of Remote Sensing,* 31(7): 1679-1697.

Botha, E., B. Zebarth, and B. Leblon, 2006. Non-destructive estimation of potato leaf chlorophyll and protein contents from hyperspectral measurements using the PROSPECT radiative transfer model. *Canadian journal of plant science,* 86(1): 279-291.

Bouman, B., 1992. Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. *International Journal of Remote Sensing,* 13(14): 2565-2581.

Broge, N.H., and E. Leblanc, 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment,* 76(2): 156-172.

Bureau, S., D. Ruiz, M. Reich, B. Gouble, D. Bertrand, J.M. Audergon, and C.M.G.C. Renard, 2009. Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy. *Food Chemistry,* 113(4): 1323-1328.

Card, D.H., Peterson, D.L., Matson, P.A. and Aber, J.D., 1988. Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. Remote Sensing of Environment, 26(2): 123-147.

Carrascal, L.M., I. Galván, and O. Gordo, 2009. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos,* 118(5): 681-690.

Ceccato, P., S. Flasse, S. Tarantola, S. Jacquemoud, and J.M. Grégoire, 2001. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment,* 77(1): 22-33.

Chang, C.I., 2003. *Hyperspectral imaging: Techniques for spectral detection and classification,* Springer, New York, 370 p.

Chen, Q., J. Zhao, X. Huang, H. Zhang, and M. Liu, 2006. Simultaneous determination of total polyphenols and caffeine contents of green tea by near-infrared reflectance spectroscopy. *Microchemical Journal,* 83(1): 42-47.

Chen, Q., J. Zhao, M. Liu, J. Cai, and J. Liu, 2008a. Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms. *Journal of Pharmaceutical and Biomedical Analysis,* 46(3): 568-573.

Chen, Q., J. Zhao, and S. Vittayapadung, 2008b. Identification of the green tea grade level using electronic tongue and pattern recognition. *Food Research International,* 41(5): 500-504.

Chen, X., and X. Lei, 2009. Application of a Hybrid Variable Selection Method for Determination of Carbohydrate Content in Soy Milk Powder Using Visible and Near Infrared Spectroscopy. *Journal of Agricultural and Food Chemistry,* 57(2): 334-340.

Cheng, Z., L. Zhang, H. Liu, and A. Zhu, 2009. Successive projections algorithm and its application to selecting the wheat near-infrared spectral variables. *Spectroscopy and Spectral Analysis,* 20(4): 949-952.

Cho, M.A., 2007. *Hyperspectral remote sensing of biochemical and biophysical parameters,* Wageingen University, Wageingen, 206 p.

Cho, M.A., and A.K. Skidmore, 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sensing of Environment,* 101(2): 181-193.

Cho, M.A., A.K. Skidmore, F. Corsi, S.E. van Wieren, and I. Sobhan, 2007. Estimation of green grass/herb biomass from airborne hyperspectral

imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation,* 9(4): 414-424.

Chong, K., T. Ho, and M. Jalloh, 2009. Soil Nitrogen Phosphorus and Tea Leaf Growth in Organic and Conventional Farming of Selected Fields at Sabah Tea Plantation Slope. *Journal of Sustainable Development,* 1(3): P117.

Chu, D.C., T. Okubo, Y. Nagato, and H. Yokogoshi, 1999. L-theanine - A unique amino acid of green tea and its relaxation effect in humans. *Trends in Food Science and Technology,* 10(6-7): 199-204.

Clark, M.L., D.A. Roberts, J.J. Ewel, and D.B. Clark, 2011. Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. *Remote Sensing of Environment*

Clevers, J., L. Kooistra, and M. Schaepman, 2010. Estimating canopy water content using hyperspectral remote sensing data. *International Journal of Applied Earth Observation and Geoinformation,* 12(2): 119-125.

Coleman, T., M.A. Branch, and A. Grace, 1999. *Optimization Toolbox for Use with MATLAB: User's Guide, Version 2,* The Math Works, Incorporated.

Combal, B., F. Baret, and M. Weiss, 2002. Improving canopy variables estimation from remote sensing data by exploiting ancillary information - case study on sugar beet canopies. *Agronomie,* 22(2): 205-215.

Combal, B., F. Baret, M. Weiss, A. Trubuil, D. Mace, A. Pragnere, R. Myneni, Y. Knyazikhin, and L. Wang, 2003. Retrieval of canopy biophysical variables from bidirectional reflectance:: Using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment,* 84(1): 1-15.

Curran, P.J., 1989. Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30(3): 271-278.

Curran, P.J., J.L. Dungan, B.A. Macler, S.E. Plummer, and D.L. Peterson, 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sensing of Environment,* 39(2): 153-166.

Curran, P.J., J.L. Dungan, and D.L. Peterson, 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment,* 76(3): 349-359.

Dabrowska-Zielinska, K., M. Gruszczynska, S. Lewinski, A. Hoscilo, and J. Bojanowski, 2009. Application of remote and in situ information to the management of wetlands in Poland. *Journal of Environmental Management,* 90(7): 2261-2269.

Dang, M.V., 2005. Soil－plant nutrient balance of tea crops in the northern mountainous region, Vietnam. *Agriculture, Ecosystems & Environment,* 105(1): 413-418.

Darvishzadeh, R., C. Atzberger, A. Skidmore, and M. Schlerf, 2011. Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66(6): 894-906.

Darvishzadeh, R., A. Skidmore, C. Atzberger, and S. van Wieren, 2008a. Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture. *International Journal of Applied Earth Observation and Geoinformation,* 10(3): 358-373.

Darvishzadeh, R., A. Skidmore, M. Schlerf, and C. Atzberger, 2008b. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment,* 112(5): 2592-2604.

Darvishzadeh, R., A. Skidmore, M. Schlerf, C. Atzberger, F. Corsi, and M. Cho, 2008c. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63(4): 409-426.

Davey, M.W., W. Saeys, E. Hof, H. Ramon, R.L. Swennen, and J. Keulemans, 2009. Application of visible and near-infrared reflectance spectroscopy (Vis/NIRS) to determine carotenoid contents in banana (*Musa spp.*) fruit pulp. *Journal of Agricultural and Food Chemistry,* 57(5): 1742-1751.

Dawson, T., 2000. The potential for estimating chlorophyll content from a vegetation canopy using the Medium Resolution Imaging Spectrometer (MERIS). *International Journal of Remote Sensing,* 21(10): 2043-2051.

De Jong, S., E. Pebesma, and B. Lacaze, 2003. Above-ground biomass assessment of Mediterranean forests using airborne imaging spectrometry: the DAIS Peyne experiment. *International Journal of Remote Sensing,* 24(7): 1505-1520.

Delalieux, S., B. Somers, S. Hereijgers, W. Verstraeten, W. Keulemans, and P. Coppin, 2008. A near-infrared narrow-waveband ratio to determine Leaf Area Index in orchards. *Remote Sensing of Environment,* 112(10): 3762-3772.

Dinar, A., and R. Mendelsohn, 2012. *Handbook on Climate Change and Agriculture,* Edward Elgar Publishing, Cheltenham, UK, 544 p.

Ding, Y., H. Yu, and S. Mou, 2002. Direct determination of free amino acids and sugars in green tea by anion-exchange chromatography with integrated pulsed amperometric detection. *Journal of Chromatography A,* 982(2): 237-244.

Draženka, K., H. Dunja, B. Ana, G.K. Kovačević, and V. Ivana, 2010. Green tea preparation and its influence on the content of bioactive compounds. *Food Research International,* 43(1): 167-176.

Duckworth, J.H., 1998. *Spectroscopic quantitative analysis, Applied Spectroscopy: A Compact Reference for Practioners* (J. Workman and

A.W. Springsteen, editors), Academic Press, San Diego and Chestnut Hill, USA, pp. 166-176.

Dutta, R. (2011). A spatio-temporal analysis of tea productivity and quality in North East India. In, *ITC Printing Department* (p. 158). Enschede: University of Twente

Dutta, R., K.R. Kashwan, M. Bhuyan, E.L. Hines, and J.W. Gardner, 2003. Electronic nose based tea quality standardization. *Neural Networks,* 16(5): 847-853.

Dutta, R., A. Stein, and R. Bhagat, 2011. Integrating satellite images and spectroscopy to measuring green and black tea quality. *Food Chemistry,* 127(2): 866-874.

Dutta, R., A. Stein, E. Smaling, R. Bhagat, and M. Hazarika, 2010. Effects of Plant Age and Environmental and Management Factors on Tea Yield in Northeast India. *Agronomy Journal,* 102(4): 1290-1301.

Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing,* 11(10): 1775-1795.

Fang, H., S. Liang, and A. Kuusk, 2003. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing of Environment,* 85(3): 257-270.

FAO (2012). World tea production. In: FAOSTAT. Retrieved at http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567

Farifteh, J., F. Van der Meer, C. Atzberger, and E.J.M. Carranza, 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment,* 110(1): 59-78.

Fava, F., R. Colombo, S. Bocchi, M. Meroni, M. Sitzia, N. Fois, and C. Zucca, 2009. Identification of hyperspectral vegetation indices for Mediterranean pasture characterization. *International Journal of Applied Earth Observation and Geoinformation,* 11(4): 233-243.

Feret, J.B., C. François, G.P. Asner, A.A. Gitelson, R.E. Martin, L.P.R. Bidel, S.L. Ustin, G. le Maire, and S. Jacquemoud, 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment,* 112(6): 3030-3043.

Ferwerda, J.G. (2005). Charting the quality of forage : measuring and mapping the variation of chemical components in foliage with hyperspectral remote sensing. In, *ITC Dissertation* (p. 166). Wageningen: Wageningen University

Ferwerda, J.G., and A.K. Skidmore, 2007. Can nutrient status of four woody plant species be predicted using field spectrometry? *ISPRS Journal of Photogrammetry and Remote Sensing,* 62(6): 406-414.

Ferwerda, J.G., A.K. Skidmore, and O. Mutanga, 2005. Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International Journal of Remote Sensing,* 26(18): 4083-4095.

Ferwerda, J.G., A.K. Skidmore, and A. Stein, 2006a. A bootstrap procedure to select hyperspectral wavebands related to tannin content. *International Journal of Remote Sensing,* 27(7): 1413 - 1424.

Ferwerda, J.G., A.K. Skidmore, S.E. Van Wieren, and H.H.T. Prins, 2006b. Soil N and P affected reflectance signatures of individual Colophospermum mopane leaves. *Journal of Spatial Science,* 51(2): 103-113.

Fourty, T., F. Baret, S. Jacquemoud, G. Schmuck, and J. Verdebout, 1996. Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. *Remote Sensing of Environment,* 56(2): 104-117.

Frei, B., and J.V. Higdon, 2003. Antioxidant activity of tea polyphenols in vivo: evidence from animal studies. *The Journal of nutrition,* 133(10): 3275S-3284S.

Galvão, L.S., A.R. Formaggio, E.G. Couto, and D.A. Roberts, 2008. Relationships between the mineralogical and chemical composition of tropical soils and topography from hyperspectral remote sensing data. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63(2): 259-271.

Gates, D.M., H.J. Keegan, J.C. Schleter, and V.R. Wiedner, 1965. Spectral properties of plants. *Applied Optics,* 411-20.

Gebbers, R., and V.I. Adamchuk, 2010. Precision agriculture and food security. *Science,* 327(5967): 828-831.

Gitelson, A., Y. Gritz, and M. Merzlyak, 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *journal of plant physiology,* 160(3): 271-282.

Gobron, N., B. Pinty, and M.M. Verstraete, 1997. Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. *Geoscience and Remote Sensing, IEEE Transactions on,* 35(6): 1438-1445.

Goel, P., S. Prasher, J. Landry, R. Patel, R. Bonnell, A. Viau, and J. Miller, 2003. Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. *Computers and Electronics in Agriculture,* 38(2): 99-124.

Gomez, C., P. Lagacherie, and G. Coulouma, 2008. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma,* 148(2): 141-148.

Graham, H.N., 1992. Green tea composition, consumption, and polyphenol chemistry. *Preventive Medicine,* 21(3): 334-350.

Green, R.O., M.L. Eastwood, C.M. Sarture, T.G. Chrien, M. Aronsson, B.J. Chippendale, J.A. Faust, B.E. Pavri, C.J. Chovit, and M. Solis, 1998. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment,* 65(3): 227-248.

Gu, Q., J. Lu, and B. Ye, 2005. *Tea Chemistry,* University of Science and Technology of China Publishing, Hefei, 381 p.

Gulati, A., S.D. Ravindranath, and A.K. Gupta, 1999. Variation in chemical composition and quality of tea (*Camellia sinensis*) with increasing blister blight (*Exobasidium vexans*) severity. *Mycological Research,* 103(11): 1380-1384.

Guo, Z., L. Chen, C. Zhao, W. Huang, and Q. Chen, 2012, Nondestructive Estimation of Total Free Amino Acid in Green Tea by Near Infrared Spectroscopy and Artificial Neural Networks. *Proceedings of the IFIP Advances in Information and Communication Technology,* Germany, pp. 43-53.

Haaland, D.M., and E.V. Thomas, 1988. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry,* 60(11): 1193-1202.

Haboudane, D., J.R. Miller, N. Tremblay, P.J. Zarco-Tejada, and L. Dextraze, 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment,* 81(2-3): 416-426.

Hacisalihoglu, G., B. Larbi, and A.M. Settles, 2009. Near-infrared reflectance spectroscopy predicts protein, starch, and seed weight in intact seeds of common bean (*Phaseolus vulgaris* L.). *Journal of Agricultural and Food Chemistry,* 58(2): 702-706.

Hansen, P., J.R. Jørgensen, and A. Thomsen, 2002. Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. *The Journal of Agricultural Science,* 139(03): 307-318.

Hansen, P.M., and J.K. Schjoerring, 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment,* 86(4): 542-553.

Hauksson, J.B., G. Bergqvist, U. Bergsten, M. Sjöström, and U. Edlund, 2001. Prediction of basic wood properties for Norway spruce. Interpretation of near infrared spectroscopy data using partial least squares regression. *Wood Science and Technology,* 35(6): 475-485.

He, K.S., D. Rocchini, M. Neteler, and H. Nagendra, 2011. Benefits of hyperspectral remote sensing for tracking plant invasions. *Diversity and Distributions,* 17(3): 381-392.

Healey, G., and D. Slater, 1999. Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions. *Geoscience and Remote Sensing, IEEE Transactions on,* 37(6): 2706-2717.

Hilton, P.J., R. Palmer-Jones, and R.T. Ellis, 1973. Effects of season and nitrogen fertiliser upon the flavanol composition and tea making quality

of fresh shoots of tea (*Camellia sinensis* L.) in central Africa. *Journal of the Science of Food and Agriculture,* 24(7): 819-826.

Houborg, R., and E. Boegh, 2008. Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data. *Remote Sensing of Environment,* 112(1): 186-202.

Huang, Z., B.J. Turner, S.J. Dury, I.R. Wallis, and W.J. Foley, 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment,* 93(1-2): 18-29.

Huber, S., M. Kneubühler, A. Psomas, K. Itten, and N.E. Zimmermann, 2008. Estimating foliar biochemistry from hyperspectral data in mixed forest canopy. *Forest Ecology and Management,* 256(3): 491-501.

Huete, A., 1989. *Soil influences in remotely sensed vegetation-canopy spectra, Theory and Applications of Optical Remote Sensing* (G. Asrar editor), Wiley, pp. 112-148

Hui, R., D. Hou, T. Li, and C. Guan, 2002. Studyies on Chemical Components of Natural Tea. *Journal of Anshan Normal University,* 4(1): 54 - 57.

Hung, Y.T., P.C. Chen, R.L.C. Chen, and T.J. Cheng, 2010. Sequential determination of tannin and total amino acid contents in tea for taste assessment by a fluorescent flow-injection analytical system. *Food Chemistry,* 118(3): 876-881.

Hussain, I., M. Saleem, Y. Iqbal, and S. Khalil, 2011. Comparison of Vitamin C contents in Commercial Tea Brands and Fresh Tea Leaves. *Journal of the Chemical Society of Pakistan,* 28(5): 421.

Iaquinta, J., B. Pinty, and J.L. Privette, 1997. Inversion of a physically based bidirectional reflectance model of vegetation. *Geoscience and Remote Sensing, IEEE Transactions on,* 35(3): 687-698.

Ishikawa, D., I. Etsuji, S. Sekioka, T. Okamoto, A. Nesumi, and K. Oba, 2006, Development of yield and quality estimation methods for tea-plant by image processing. *Proceedings of the ASA-CSSA-SSSA 2006 International Meetings,* Nov. 12-16, Indianapolis, pp.

Islam, S., and R. Kothari, 2000. Artificial neural networks in remote sensing of hydrologic processes. *Journal of Hydrologic Engineering,* 5(2): 138-144.

Iwasa, K., and H. Torii, 1962. A colorimetric determination of tea tannin with ferrous tartrate. *Study of Tea,* 26(1): 87-91.

Jacquemoud, S., C. Bacour, H. Poilve, and J.P. Frangi, 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing of Environment,* 74(3): 471-481.

Jacquemoud, S., and F. Baret, 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment,* 34(2): 75-91.

Jacquemoud, S., S. Ustin, J. Verdebout, G. Schmuck, G. Andreoli, and B. Hosgood, 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment,* 56(3): 194-202.

Jacquemoud, S., W. Verhoef, F. Baret, C. Bacour, P.J. Zarco-Tejada, G.P. Asner, C. François, and S.L. Ustin, 2009. PROSPECT+ SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment,* 113S56-S66.

Joffre, R., D. Gillon, P. Dardenne, R. Agneessens, and R. Biston, 1992. The use of near-infrared reflectance spectroscopy in litter decomposition studies. *Annals of Forest Science,* 49(5): 481-488.

Juneja, L.R., D.C. Chu, T. Okubo, Y. Nagato, and H. Yokogoshi, 1999. L-theanine--a unique amino acid of green tea and its relaxation effect in humans. *Trends in Food Science & Technology,* 10(6-7): 199-204.

Kaneko, S., K. Kumazawa, H. Masuda, A. Henze, and T. Hofmann, 2006. Molecular and sensory studies on the umami taste of Japanese green tea. *Journal of Agricultural and Food Chemistry,* 54(7): 2688-2694.

Khanchi, A.R., M.K. Mahani, M. Hajihosseini, M.G. Maragheh, M. Chaloosi, and F. Bani, 2007. Simultaneous spectrophotometric determination of caffeine and theobromine in Iranian tea by artificial neural networks and its comparison with PLS. *Food Chemistry,* 103(3): 1062-1068.

Knox, N.M., A.K. Skidmore, H.H.T. Prins, G.P. Asner, H. van der Werff, W.F. de Boer, C. van der Waal, H.J. de Knegt, E.M. Kohi, and R. Slotow, 2011. Dry season mapping of savanna forage quality, using the hyperspectral Carnegie Airborne Observatory sensor. *Remote Sensing of Environment,* 115(6): 1478-1488

Kobayashi, K., Y. Nagato, N. Aoi, L. Juneja, M. Kim, T. Yamamoto, and S. Sugimoto, 1998. Effects of L-theanine on the release of alpha-brain waves in human volunteers. *Nippon Nogeikagaku Kaishi,* 72(2): 153-157.

Kokaly, R.F., G.P. Asner, S.V. Ollinger, M.E. Martin, and C.A. Wessman, 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment,* 113(SUPPL. 1): S78-S91.

Kokaly, R.F., and R.N. Clark, 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment,* 67(3): 267-287.

Kooistra, L., E.A.L. Salas, J.G.P.W. Clevers, R. Wehrens, R.S.E.W. Leuven, P.H. Nienhuis, and L.M.C. Buydens, 2004. Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains. *Environmental Pollution,* 127(2): 281-290.

Kumar, L., K.S. Schmidt and A.K. Skidmore, 2001. *Imaging spectroscopy and vegetation science*, *Imaging Spectrometry* (S.M.D. Jong editor), Academic Kluwer Press, Dordrecht, The Netherlands, pp. 111-155.

Kumar, N.S., and M. Rajapaksha, 2005. Separation of catechin constituents from five tea cultivars using high-speed counter-current chromatography. *Journal of Chromatography A,* 1083(1): 223-228.

Kutner, M.H., C.J. Nachtsheim, and J. Neter, 2004. *Applied Linear Statistical Models,* McGraw-Hill/Irwin, 1396 p.

Le Maire, G., C. Francois, and E. Dufrene, 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment,* 89(1): 1-28.

Lee, K.S., M.J. Kook, J.I. Shin, S.H. Kim, and T.G. Kim, 2007. Spectral characteristics of forest vegetation in moderate drought condition observed by laboratory measurements and spaceborne hyperspectral data. *Photogrammetric engineering and remote sensing,* 73(10): 1121-1127.

Lee, Y.P., and T. Takahashi, 1966. An improved colorimetric determination of amino acids with the use of ninhydrin. *Analytical Biochemistry,* 14(1): 71-77.

Lee, Z., K.L. Carder, S.K. Hawes, R.G. Steward, T.G. Peacock, and C.O. Davis, 1994. Model for the interpretation of hyperspectral remote-sensing reflectance. *Applied Optics,* 33(24): 5721-5732.

Lera, G., and M. Pinzolas, 2002. Neighborhood based Levenberg-Marquardt algorithm for neural network training. *IEEE Transactions on Neural Networks,* 13(5): 1200-1203.

Li, X., and Y. He, 2008. Discriminating varieties of tea plant based on Vis/NIR spectral characteristics and using artificial neural networks. *Biosystems Engineering,* 99(3): 313-321.

Li, X., and A.H. Strahler, 1992. Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Transactions on geoscience and remote sensing,* 30(2): 276-292.

Li, Z. (2002). Studies on the dynamic changes of the main quality components in green tea brewing and the correlation between the components and the organoleptic evaluation. In, *Department of Tea Science, Zhejiang University* (p. 51). Hangzhou: Zhejiang University

Liang, C., Y. Yajun, Y. Fulian, G. Qikang, and C. Daming, 1998. Genetic Diversity of 15 Tea (*Camellia sinensis* (L.) O. Kuntze) Cultivars Using RAPD Markers. *Journal of Tea Science,* 18(1): 21-27

Liang, H., and X. Liu, 2008, Hyperspectral analysis of leaf copper accumulation in agronomic crop based on artificial neural network. *Proceedings of the 2008 International Workshop on Earth Observation and Remote Sensing Applications, EORSA,* pp. 1-6

Liang, S., 2004. *Quantitative remote sensing of land surfaces,* Wiley-IEEE, 534 p.

Liang, Y., J. Lu, L. Zhang, S. Wu, and Y. Wu, 2003. Estimation of black tea quality by analysis of chemical composition and colour difference of tea infusions. *Food Chemistry,* 80(2): 283-290.

Liang, Y.R., Q. Ye, J. Jin, H. Liang, J.L. Lu, Y.Y. Du, and J.J. Dong, 2008. Chemical and instrumental assessment of green tea sensory preference. *International Journal of Food Properties,* 11(2): 258-272.

Lin, G.F., and L.H. Chen, 2004. A non-linear rainfall-runoff model using radial basis function network. *Journal of Hydrology,* 289(1-4): 1-8.

Lin, Y.L., I.M. Juan, Y.L. Chen, Y.C. Liang, and J.K. Lin, 1996. Composition of polyphenols in fresh tea leaves and associations of their oxygen-radical-absorbing capacity with antiproliferative actions in fibroblast cells. *Journal of Agricultural and Food Chemistry,* 44(6): 1387-1394.

Lin, Z.H., Y.P. Qi, R.B. Chen, F.Z. Zhang, and L.S. Chen, 2012. Effects of phosphorus supply on the quality of green tea. *Food Chemistry,* 130(4): 908-914.

Liu, F., Z.L. Jin, M.S. Naeem, T. Tian, F. Zhang, Y. He, H. Fang, Q.F. Ye, and W.J. Zhou, 2010. Applying Near-Infrared Spectroscopy and Chemometrics to Determine Total Amino Acids in Herbicide-Stressed Oilseed Rape Leaves. *Food and Bioprocess Technology,* 4(7): 1-8.

Liu, T., and L. Yue, 2005. *Chinese tea,* China Intercontinental Press, Beijing, 112 p.

Lu, J., F. We, and C. Li, 1994. A study on the relationship between main chemical components and quality of tea varieties. *Southwest China Journal of Agricultureal Sciences,* 7(Sup.): 1-4.

Lu, Y., H. Chen, J. Lv, and X. Chen, 2005. Near infrared determination of catechin in tea polyphenol. *Chinese Journal of Analytical Chemistry,* 33(6): 835-837.

Luypaert, J., M.H. Zhang, and D.L. Massart, 2003. Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea, *Camellia sinensis* (L.). *Analytica Chimica Acta,* 478(2): 303-312.

Maeda, A.S., S. Buzetti, and A.C. Bolonhezi, 2009. Effects of nitrogen and potassium fertilizers on quality and yield of sugarcane crop. *STAB - Açúcar, Álcool e Subprodutos,* 27(4): 40, 42-45.

Majeke, B., J.A.N. Van Aardt, and M.A. Cho, 2008. Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests,* 70(3): 275-285.

Marcoulides, G.A., and C. Saunders, 2006. PLS: A silver bullet? *MIS Quarterly: Management Information Systems,* 30(2): III-IV.

Martin, L.C., 2007. *Tea: the drink that changed the world,* Tuttle Publishing, North Clarendon, 256 p.

Martin, M.E., and J.D. Aber, 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications,* 7(2): 431-443.

Mas, J.F., and J.J. Flores, 2008. The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing,* 29(3): 617-663.

McKay, D.L., and J.B. Blumberg, 2002. The role of tea in human health: An update. *Journal of the American College of Nutrition,* 21(1): 1-13.

Miao, Q., W. Zhao, X. Guo, K. Liu, J. Han, and Z. Wang, 2011, Inversion of reed leaf chlorophyll content based on PROSPECT model. *Proceedings of the 19th International Conference on Geoinformatics,* pp. 1-4.

Mitscher, L.A., and V. Dolby, 1997. *The Green Tea Book: China's Fountain of Youth,* Avery, p.

Moore, B., 1981. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control,* 26(1): 17-32.

Moore, D.S., and G.P. McCabe, 1998. *Introduction to the practice of statistics,* W. H. Freeman and Company, New York, 896 p.

Moran, M.S., Y. Inoue, and E. Barnes, 1997. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing of Environment,* 61(3): 319-346.

Mukhtar, H., and N. Ahmad, 2000. Tea polyphenols: prevention of cancer and optimizing health. *The Americal journal of Clinical Nutrition,* 71(6): 1698-1702.

Mutanga, O., and R. Ismail, 2010. Variation in foliar water content and hyperspectral reflectance of Pinus patula trees infested by Sirex noctilio. *Southern Forests,* 72(1): 1-7.

Mutanga, O., and L. Kumar, 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing,* 28(21): 4897-4911.

Mutanga, O., and A.K. Skidmore, 2004a. Hyperspectral band depth analysis for a better estimation of grass biomass (*Cenchrus ciliaris*) measured under controlled laboratory conditions. *International Journal of Applied Earth Observation and Geoinformation,* 5(2): 87-96.

Mutanga, O., and A.K. Skidmore, 2004b. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment,* 90(1): 104-115.

Mutanga, O., A.K. Skidmore, L. Kumar, and J. Ferwerda, 2005. Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing,* 26(6): 1093-1108.

Mutanga, O., A.K. Skidmore, and H.H.T. Prins, 2004a. Discriminating sodium concentration in a mixed grass species environment of the Kruger National Park using field spectrometry. *International Journal of Remote Sensing,* 25(20): 4191-4201.

Mutanga, O., A.K. Skidmore, and H.H.T. Prins, 2004b. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment,* 89(3): 393-408.

Mutanga, O., J. Van Aardt, and L. Kumar, 2009. Imaging spectroscopy (hyperspectral remote sensing) in southern Africa: an overview. *South African Journal of Science,* 105(5 & 6): 193-198.

Nakagawa, M., 1970. Constituents in tea leaf and their contribution to the taste of green tea liquor. *Japan Agricultural Research Quarterly,* 543-47.

Nakagawa, M., 1975. Chemical components and taste of green tea. *Japan Agricultural Research Quarterly,* 9(3): 156-160.

Nicolaï, B.M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, and J. Lammertyn, 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology,* 46(2): 99-118.

Norris, K.H., 1964. Design and development for a new moisture meter. *Agric. Eng.,* 45370-372.

Norris, K.H., R.F. Barnes, J.E. Moore, and J.S. Shenk, 1976. Predicting Forage Quality by Infrared Replectance Spectroscopy. *Journal of Animal Science,* 43(4): 839-878.

Obanda, M., P. Okinda Owuor, and R. Mang'oka, 2001. Changes in the chemical and sensory quality parameters of black tea due to variations of fermentation time and temperature. *Food Chemistry,* 75(4): 395-404.

Obanda, M., P.O. Owuor, and S.J. Taylor, 1997. Flavanol composition and caffeine content of green leaf as quality potential indicators of Kenyan black teas. *Journal of the Science of Food and Agriculture,* 74(2): 209-215.

Ohnsmann, J., G. Quintás, S. Garrigues, and M. de la Guardia, 2002. Determination of caffeine in tea samples by Fourier transform infrared spectrometry. *Analytical and bioanalytical chemistry,* 374(3): 561-565.

Osborne, B., and T. Fearn, 1986. *Near infrared spectroscopy in food analysis,* Longman scientific and technical, London, 200 p.

Ott, L., and M. Longnecker, 2010. *An introduction to statistical methods and data analysis,* Cengage Learning, Belmont, 1273 p.

Owuor, P.O., D.M. Kamau, S.M. Kamunya, S.W. Msomba, M.A. Uwimana, A.W. Okal, and B.O. Kwach, 2011. *Effects of Genotype, Environment and Management on Yields and Quality of Black Tea, Genetics, Biofuels and Local Farming Systems* (E. Lichtfouse editor), Springer Netherlands, pp. 277-307.

Owuor, P.O., and M. Obanda, 1998. The changes in black tea quality due to variations of plucking standard and fermentation time. *Food Chemistry,* 61(4): 435-441.

Owuor, P.O., and M. Obanda, 2007. The use of greentea (*Camellia sinensis*) leaf flavan-3-ol composition in predicting plain black tea quality potential. *Food Chemistry,* 100(3): 873-884.

Peng, J., Q. Liu, J. Li, and Q. Liu, 2012. *Quantitative Analysis of Crops Chlorophyll in the Heihe River Basin by Hyperspectral Remote Sensing Image, Recent Advances in Computer Science and Information*

Engineering (Z. Qian, L. Cao, W. Su, T. Wang and H. Yang, editors), Springer Berlin Heidelberg, Berlin, pp. 681-689.

Peng, Y., W. Wang, H. Huang, X. Wang, and X. Gao, 2009, Prediction of chlorophyll content of winter wheat using leaf-level hyperspectral imaging data. *Proceedings of the American Society of Agricultural and Biological Engineers Annual International Meeting 2009,* pp. 3756-3766.

Penuelas, J., J. Gamon, A. Fredeen, J. Merino, and C. Field, 1994. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment,* 48(2): 135-146.

Peterson, D.L., J.D. Aber, P.A. Matson, D.H. Card, N. Swanberg, C. Wessman, and M. Spanner, 1988. Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment,* 24(1): 85-108,IN101-IN119.

Pinter, P.J., J.L. Hatfield, J.S. Schepers, E.M. Barnes, M.S. Moran, C.S.T. Daughtry, and D.R. Upchurch, 2003. Remote sensing for crop management. *Photogrammetric engineering and remote sensing,* 69(6): 647-664.

Pojić, M., J. Mastilović, and N. Majcen, 2012. Robustness of the near infrared spectroscopy method determined using univariate and multivariate approach. *Food Chemistry,* 134(3): 1699-1705.

Qin, W., and S. Liang, 2000. Plane - parallel canopy radiation transfer modeling: Recent advances and future directions. *Remote Sensing Reviews,* 18(2-4): 281-305.

Rajapakse, R., 2002. Spectral characterization and LAI modelling for the tea (*Camellia sinensis* (L.) O. Kuntze) canopy. *International Journal of Remote Sensing,* 23(18): 3569-3577.

Ramoelo, A., M. Cho, R. Mathieu, A. Skidmore, M. Schlerf, I. Heitkonig, and H. Prins (2011a). Integrating environmental and in situ hyperspectral remote sensing variables for grass nitrogen estimation in savannah ecosystems. In, *The International Symposium on Remote Sensing of Environment (ISRSE 2011)-34th ISRSE*. Sydney, Australia: 10-15 April 2011

Ramoelo, A., A.K. Skidmore, M. Schlerf, R. Mathieu, and I. Heitkönig, 2011b. Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66(4): 408-417.

Rao, N.R., M. Kapoor, N. Sharma, and K. Venkateswarlu, 2007. Yield prediction and waterlogging assessment for tea plantation land using satellite image-based techniques. *International Journal of Remote Sensing,* 28(7): 1561 - 1576.

Ravichandran, R., and R. Parthiban, 2000. Lipid occurrence, distribution and degradation to flavour volatiles during tea processing. *Food Chemistry,* 68(1): 7-13.

Rossel, R.A.V., 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems,* 90(1): 72-83.

Ruan, J., R. Haerdter, and J. Gerendás, 2010. Impact of nitrogen supply on carbon/nitrogen allocation: a case study on amino acids and catechins in green tea [*Camellia sinensis* (L.) O. Kuntze] plants. *Plant Biology,* 12(5): 724-734.

Sáiz-Abajo, M., J. González-Sáiz, and C. Pizarro, 2006. Prediction of organic acids and other quality parameters of wine vinegar by near-infrared spectroscopy. A feasibility study. *Food Chemistry,* 99(3): 615-621.

Sandmeier, S., C. Müller, B. Hosgood, and G. Andreoli, 1998. Physical mechanisms in hyperspectral BRDF data of grass and watercress. *Remote Sensing of Environment,* 66(2): 222-233.

Satyanarayana, N., and V. Sharma, 1986. Tea (*Camellia* L. Sp.) Germplasm in South India. *Plantation Crops,* 2173-179.

Savitzky, A., and M.J.E. Golay, 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry,* 36(8): 1627-1639.

Scharbert, S., and T. Hofmann, 2005. Molecular definition of black tea taste by means of quantitative studies, taste reconstitution, and omission experiments. *Journal of Agricultural and Food Chemistry,* 53(13): 5377-5384.

Scharbert, S., N. Holzmann, and T. Hofmann, 2004. Identification of the astringent taste compounds in black tea infusions by combining instrumental analysis and human bioresponse. *Journal of Agricultural and Food Chemistry,* 52(11): 3498-3508.

Schlerf, M., and C. Atzberger, 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment,* 100(3): 281-294.

Schlerf, M., C. Atzberger, and J. Hill, 2005. Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment,* 95(2): 177-194.

Schlerf, M., C. Atzberger, J. Hill, H. Buddenbaum, W. Werner, and G. Schüler, 2010. Retrieval of chlorophyll and nitrogen in Norway spruce (*Picea abies L. Karst.*) using imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation,* 12(1): 17-26.

Schmidt, K.S., and A.K. Skidmore, 2003. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment,* 85(1): 92-108.

Schulz, H., U.H. Engelhardt, A. Wegent, H.H. Drews, and S. Lapczynski, 1999. Application of near-infrared reflectance spectroscopy to the

simultaneous prediction of alkaloids and phenolic substances in green tea leaves. *Journal of Agricultural and Food Chemistry,* 47(12): 5064-5067.

Shankar, R.R., M. Jayathilaka, and W. Wanigasundera, 2003. Assessment of manual and shear harvesting in high elevation tea plantations in Sri Lanka. *Tropical Agricultural Research,* 1598-108.

Sharangi, A.B., 2009. Medicinal and therapeutic potentialities of tea (*Camellia sinensis* L.) - A review. *Food Research International,* 42(5-6): 529-535.

Si, Y., M. Schlerf, R. Zurita-Milla, A. Skidmore, and T. Wang, 2012. Mapping spatio-temporal variation of grassland quantity and quality using MERIS data and the PROSAIL model. *Remote Sensing of Environment,* 121415-425.

Sin, H.L., 2007. Green tea consumption and mortality in Japan. *Journal of the American Medical Association,* 297(4): 360-361.

Singh, H.P., S. Ravindranath, and C. Singh, 1999. Analysis of tea shoot catechins: spectrophotometric quantitation and selective visualization on two-dimensional paper chromatograms using diazotized sulfanilamide. *Journal of Agricultural and Food Chemistry,* 47(3): 1041-1045.

Singhal, R.S., P.R. Kulkarni, and D.V. Rege, 1997. *Handbook of indices of food quality and authenticity,* Woodhead Publishing, Abington, 561 p.

Skidmore, A.K., J.G. Ferwerda, O. Mutanga, S.E. Van Wieren, M. Peel, R.C. Grant, H.H.T. Prins, F.B. Balcik, and V. Venus, 2010. Forage quality of savannas - Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sensing of Environment,* 114(1): 64-72.

Skidmore, A.K., B.J. Turner, W. Brinkhof, and E. Knowles, 1997. Performance of a neural network: Mapping forests using GIS and remotely sensed data. *Photogrammetric engineering and remote sensing,* 63(5): 501-514.

Song, S., W. Gong, B. Zhu, and X. Huang, 2011. Wavelength selection and spectral discrimination for paddy rice, with laboratory measurements of hyperspectral leaf reflectance. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66(5): 672-682.

Soukupova, J., B.N. Rock, and J. Albrechtova, 2002. Spectral characteristics of lignin and soluble phenolics in the near infrared - a comparative study. *International Journal of Remote Sensing,* 23(15): 3039 - 3055.

Stagakis, S., N. Markos, O. Sykioti, and A. Kyparissis, 2010. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment,* 114(5): 977-994.

Stone, H., and J.L. Sidel, 2004. *Sensory evaluation practices,* Elsevier Academic press, London, 408 p.

Stuckens, J., W.W. Verstraeten, S. Delalieux, R. Swennen, and P. Coppin, 2009. A dorsiventral leaf radiative transfer model: Development,

validation and improved model inversion techniques. *Remote Sensing of Environment,* 113(12): 2560-2573.

Sumpio, B.E., A.C. Cordova, D.W. Berke-Schlessel, F. Qin, and Q.H. Chen, 2006. Green tea, the ‚ÄúAsian paradox,‚Äù and cardiovascular disease. *Journal of the American College of Surgeons,* 202(5): 813-825.

Suratgar, A.A., Tavakoli, M.B. and Hoseinabadi, A., 2005. Modified Levenberg-Marquardt method for neural networks training. *World academy of science, Engineering and Technology*, 6:46-48.

Tang, S., J. Chen, Q. Zhu, X. Li, M. Chen, R. Sun, Y. Zhou, F. Deng, and D. Xie, 2006. LAI inversion algorithm based on directional reflectance kernels. *Journal of Environmental Management,* 85(3): 638-648.

Taylor, S., D. Baker, P. Owuor, J. Orchard, C. Othieno, and C. Gay, 1992. A model for predicting black tea quality from the carotenoid and chlorophyll composition of fresh green tea leaf. *Journal of the Science of Food and Agriculture,* 58(2): 185-191.

Teranishi, R., E.L. Wick, and r. HornsteinI (Eds.) (1999). *Flavour chemistry: thirty years of progress*. New York: Springer

Thanh Mai, N.T., T.B. Tram, and T.T. Chien, 2012. Study on setting up the standard for tea leaves in processing of green tea powder of high quality. *Vietnam Journal of Biology,* 34(2): 224-227.

Thenkabail, P.S., R.B. Smith, and E. De Pauw, 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment,* 71(2): 158-182.

Thomas, J., A.K.A. Mandal, R.R. Kumar, and A. Chordia, 2009. Role of Biologically Active Amino Acid Formulations on Quality and Crop Productivity of Tea (*Camellia* sp.). *International Journal of Agricultural Research,* 4(7): 228-236.

Togari, N., A. Kobayashi, and T. Aishima, 1995. Relating sensory properties of tea aroma to gas chromatographic data by chemometric calibration methods. *Food Research International,* 28(5): 485-493.

Treitz, P., V. Thomas, P.J. Zarco-Tejada, P. Gong, and P.J. Curran, 2010. *Hyperspectral remote sensing for forestry,* American Society for Photogrammetry and Remote Sensing (ASPRS), 107 p.

van der Meer, F., 2004. Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation,* 5(1): 55-68.

Venkatesan, S., S. Murugesan, V.K. Senthur Pandian, and M.N.K. Ganapathy, 2005. Impact of sources and doses of potassium on biochemical and greenleaf parameters of tea. *Food Chemistry,* 90(4): 535-539.

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment,* 16(2): 125-141.

Verhoef, W., and H. Bach, 2007. Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular

surface reflectance and TOA radiance data. *Remote Sensing of Environment,* 109(2): 166-182.

Wörth, C.C.T., M. Wießler, and O.J. Schmitz, 2000. Analysis of catechins and caffeine in tea extracts by micellar electrokinetic chromatography. *Electrophoresis,* 21(17): 3634-3638.

Wang, H., and K. Helliwell, 2001. Determination of flavonols in green and black tea leaves and green tea infusions by high-performance liquid chromatography. *Food Research International,* 34(2-3): 223-227.

Wang, Y., and C.T. Ho, 2009. Polyphenolic chemistry of tea and coffee: a century of progress. *Journal of Agricultural and Food Chemistry,* 57(18): 8109-8114.

Wang, Y., F. Wang, J. Huang, X. Wang, and Z. Liu, 2009. Validation of artificial neural network techniques in the estimation of nitrogen concentration in rape using canopy hyperspectral reflectance data. *International Journal of Remote Sensing,* 30(17): 4493-4505.

Wei, J., Y. Tian, Y. Liang, and G. Zhou, 2003. Effect of Produce Autotoxic - chemical of Tea Tree on Environment Factor. *Chinese Agricultural Science Bulletin,* 19(2): 53-56.

Wei, K., L. Wang, J. Zhou, W. He, J. Zeng, Y. Jiang, and H. Cheng, 2011. Catechin contents in tea (*Camellia sinensis*) as affected by cultivar and environment and their relation to chlorophyll contents. *Food Chemistry,* 125(1): 44-48.

Wen, C.F., A.W. Dong, G.Z. Li, S. L, and Y. Lei, 2005. Determination of Total Sugar and Reducing Sugar in Viola Philippicassp Munda W. Becker by Anthrone Colorimetry. *Modern Food Science and Technology,* 21(3): 122-130.

Wessman, C.A., J.D. Aber, and D.L. Peterson, 1989. An evaluation of imaging spectrometry for estimating forest canopy chemistry. *International Journal of Remote Sensing,* 10(8): 1293-1316.

Wessman, C.A., Aber, J.D., Peterson, D.L. and Melillo, J.M., 1988a. Foliar analysis using near infrared reflectance spectroscopy. *Canadian Journal of Forest Research,* 18(1): 6-11.

Wessman, C.A., Aber, J.D., Peterson, D.L. and Melillo, J.M., 1988b. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature*, 335(6186): 154-156.

Weyer, L., 1985. Near-infrared spectroscopy of organic substances. *Applied Spectroscopy Reviews,* 21(1): 1-43.

Williams, P., F.J. El-Haramein, H. Nakkoul, and S. Rihawi, 1988. *Crop quality evaluation methods and guidelines,* International Center for Agricultural Research in the Dry Areas (ICARDA), New York, 145 p.

Williams, P., and K. Norris, 1987. *Near-infrared technology in the agricultural and food industries,* American Association of Cereal Chemists, St. Paul, MN, 330 p.

Willson, K.C., and M.N. Clifford, 1992. *Tea: Cultivation to Consumption,* Chapman & Hall, 769 p.

Wold, S., M. Sjostrom, and L. Eriksson, 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems,* 58(2): 109-130.

Workman, J., and A.W. Springsteen, 1998. *Applied Spectroscopy: A Compact Reference for Practioners,* Academic Press, San Diego etc., 539 p.

Wright, L.P., N.I.K. Mphangwe, H.E. Nyirenda, and Z. Apostolides, 2000. Analysis of caffeine and flavan-3-ol composition in the fresh leaf of *Camellia sinensis* for predicting the quality of the black tea produced in Central and Southern Africa. *Journal of the Science of Food and Agriculture,* 80(13): 1823-1830.

Xia, J., J. Li, Y. Gong, and D. Wu, 2005. Modelling for the Optimization of Fertilization of Tea High Yield. *Journal of Tea Science,* 25(3): 165-171.

Xu, H., R. Xiao, and T. Song, 2008. Effects of different fertilization on microbial biomass carbon from the red soil in tea garden. *Frontiers of Agriculture in China,* 2(4): 418-422.

Yamamoto, T., L.R. Juneja, D.C. Chu, and M. Kim, 1997. *Chemistry and Applications of Green Tea,* CRC Press, Boca Raton, 176 p.

Yan, S.H., 2007. NIR evaluation of the quality of tea and its market price. *Spectroscopy Europe,* 19(2): 16-19.

Yang, C.S., and Z.Y. Wang, 1993. Tea and cancer. *Journal of the National Cancer Institute,* 85(13): 1038-1049.

Yao, L., X. Liu, Y. Jiang, N. Caffin, B. D'Arcy, R. Singanusong, N. Datta, and Y. Xu, 2006. Compositional analysis of teas from Australian supermarkets. *Food Chemistry,* 94(1): 115-122.

Yao, M., L. Chen, and Y. Liang, 2008. Genetic diversity among tea cultivars from China, Japan and Kenya revealed by ISSR markers and its implication for parental selection in tea breeding programmes. *Plant Breeding,* 127(2): 166-172.

Yoder, B.J., and R.E. Pettigrew-Crosby, 1995. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. *Remote Sensing of Environment,* 53(3): 199-211.

Yuan, X., 2003. *Tea biochemistry (in Chinese),* China Agriculture Press, Beijing, 451 p.

Yue, Y., K. Wang, B. Zhang, and Z. Chen, 2008. Applications of Hyperspectral Remote Sensing in Ecosystem: A Review. *Remote Sensing Technology and Application,* 23(4): 471-478.

Zarco-Tejada, P.J., J.R. Miller, A. Morales, A. Berjón, and J. Agüera, 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment,* 90(4): 463-476.

Zhang, Z., and K. Friedrich, 2003. Artificial neural networks applied to polymer composites: A review. *Composites Science and Technology,* 63(14): 2029-2044.

AUTHOR'S BIOGRAPHY

Bian Meng was born on 11th December, 1981 in Shandong, China.

In 2004 she obtained her bachelor degree in Land Resources Management at East China Institute of Technology after four years studies. Then, she went to Wuhan University and enrolled in a joint master's program between Wuhan University and the International Institute for Geo-information Science and Earth Observation (ITC) in the Netherlands from 2004 to 2006. In March 2006, she completed her thesis and received her M. Sc. degree in Geo-information for Natural Resources and Environment Management. In the same year, she was awarded a doctorial scholarship from ITC, with partial financial support from Wuhan University. In 2007, she moved to the Netherlands again to start her Ph.D. research at ITC, University of Twente. Her major research interests include quantitative remote sensing, remote sensing of vegetation, crop yield and quality estimation and food quality control.

SUMMARY

Tea (*Camellia sinensis*) is an important cash crop in many countries, especially in developing countries. The quality of made tea is reflected by its smell and taste, which is traditionally appraised by professional tea tasters. This method is expensive in terms of time and labor and may also be subjective. In addition, this method can only be applied after tea leaves have been harvested and processed. Therefore, it cannot meet the requirement of quality control of tea leaves as the raw materials of tea product in the field. Although some studies have investigated the biochemical parameters in fresh tea leaves by sample surveys and traditional laboratory work, and link them to the quality of tea product, these methods are difficult to implement in a fast, non-destructive way and in a large area. With the advancement of remote sensing, hyperspectral techniques are now capable of detecting biophysical and biochemical parameters of plant respond to their local environment in space and time. Hyperspectral remote sensing has been proved to be a powerful tool in fast and non-destructive vegetation quality measurement and monitoring.

This thesis assesses the biochemical parameters of fresh tea leaves associated with tea's quality using hyperspectral remote sensing approaches at different scales (powder, leaf, canopy), evaluates the accuracy, and further explores the possibility of constructing a physically based model for tea quality estimation, aiming to study the potential of hyperspectral remote sensing for assessing the quality of tea at large scale in the future. The principal results obtained can be summarized as follows:

(1) The concentrations of biochemicals as quality indicators of tea can be estimated by means of hyperspectral remote sensing techniques, at powder, leaf and especially canopy levels.

(2) From tea powders via fresh tea leaves to tea canopies, the identified important spectral wavebands for the prediction do not overlap much. The wavebands related to the absorption features of chemicals of interest are of more importance at powder level, while relatively less important at canopy level.

(3) In descending order, we list the accuracy of the estimation of tea quality-related biochemicals at the three levels as: powder (highest), canopy (medium), fresh leaf (lowest). The prediction results at the canopy level slightly outperformed the results from leaf level, giving rise to positive expectations from airborne and space-borne sensors for mapping of tea quality, at landscape or regional scales.

(4) An improved physically based leaf radiative transfer model can help estimate the tea-polyphenols, which give the astringe taste of tea, across different cultivars and phenological stages of tea.

Overall, this study has shown the possibility of hyperspectral remote sensing techniques to retrieve tea quality parameters, by statistical modelling approaches, or by analysing the mechanism between light and layers of tea leaf chemical composition. It is anticipated that these results provide a basis for the future research, and necessary information as well as refined approaches for assessing the quality of tea using air or space-borne remote sensing over large areas. The findings contribute to the tea quality assessment before plucking. The methods proposed in this thesis can be used to help improve the management of tea garden in the further study.

SAMENVATTING

Thee (*Camellia sinensis*) is voor veel ontwikkelingslanden landen een belangrijk export gewas. De kwaliteit van thee wordt bepaald door geur en smaak die van oudsher wordt beoordeeld door professionele thee proevers. Deze methode van kwaliteit bepalen is duur, tijdrovend en mogelijk subjectief. Bovendien kan deze methode alleen worden toegepast nadat de theebladeren zijn geoogst en verwerkt. Hierdoor vind er geen kwaliteitscontrole plaats van de thee struiken in het veld. Enkele studies rapporteren over onderzoek van thee kwaliteit door middel van destructieve laboratorium metingen van verse thee bladeren. Toegepaste methodes zijn echter moeilijk te implementeren, zijn ook tijdrovend, en kunnen slechts steekproef-gewijs in grotere thee productie gebieden plaatsvinden.

Met de vooruitgang van remote sensing, oftewel ruimtelijke aardobservatie-technieken, is het nu mogelijk om met hyperspectrale technieken op te sporen hoe biofysische en biochemische parameters van planten reageren op zowel ruimte en tijd in hun lokale omgeving. Hyperspectrale remote sensing is een inmiddels bewezen krachtig instrument voor het meten en bewaken van de kwaliteit die snel is en die de vegetatie niet aantast.

Dit proefschrift evalueert en bestudeert de biochemische parameters van verse theebladeren in vergelijking met de kwaliteit van thee met gebruik van hyperspectrale remote sensing op verschillende schalen (poeder, blad, en de gehele struik). Ook wordt hiermee de juistheid bepaald en wordt een verdere mogelijkheid onderzocht voor het maken van een fysiek model die de kwaliteit van thee kan bepalen. Deze studie is gericht op de mogelijkheden die hyperspectrale remote sensing biedt om in de toekomst de kwaliteit van thee op grote schaal te beoordelen. De belangrijkste resultaten kunnen als volgt worden samengevat:

1) De concentraties van biochemische producten, als kwaliteitsindicatoren van thee, kunnen worden geschat met behulp van hyperspectrale remote sensing technieken, in het poeder, het blad en in het bijzonder de gehele theestruik.

2) Zowel van de poeder, de verse bladeren en de gehele struik, overlappen de geïdentificeerde spectrale frequentiebanden die werden voorspeld niet veel. De golflengtes die verband houden met de absorptie eigenschappen van chemische stoffen zijn van meer belang op het niveau van de poeder, terwijl het relatief minder belangrijk is op struik-niveau.

3) Als we de biochemische concentraties voorspellen, is de nauwkeurigheid het hoogst wanneer in poedervorm, gemiddeld bij de gehele thee-struik en is de nauwkeurigheid het laagst bij de (verse) bladeren. De voorspelde resultaten van de metingen op het niveau van de struiken waren nauwkeuriger dan de resultaten op het niveau van de bladeren, waardoor er positieve metingen verwacht worden van de lucht- en ruimte sensoren, die gebruikt kunnen worden voor het in kaart brengen van de kwaliteit van thee, op verschillende schalen.

4) Een fysiek verbeterde blad-stralingsoverdracht model kan de verschillende thee telers helpen om de thee polyfenolen, die de smaak van de thee maken, aangeven in de verschillende fenologische stadia.

In het algemeen heeft deze studie de mogelijkheid aangetoond om door middel van hyperspectrale remote sensing technieken parameters te verkrijgen die de kwaliteit van thee meten. Hiervoor werden statistische modelmatige benaderingen gebruikt en werd het mechanisme van licht en de verschillende lagen van de chemische compositie van theebladeren geanalyseerd. Verwacht wordt dat deze resultaten een basis zullen zijn voor toekomstig onderzoek en de nodige informatie geven over de verfijnde methoden die gebruikt zijn voor de beoordeling van de kwaliteit van thee met behulp van lucht of in de ruimte gestationeerde aardobseratie technieken over grote gebieden. Onze bevindingen kunnen op deze manier bijdragen aan de kwaliteitsbeoordeling van thee nog voor het plukken. De voorgestelde methoden in dit proefschrift kunnen worden gebruikt ter verbetering van het beheer van thee plantages in een verdere studie.

ITC DISSERTATION LIST

http://www.itc.nl/Pub/research programme/Graduate-programme/Graduate-programme-Phd_Graduates.html