



NLR-TP-98025

Theory of radiative transfer models applied in optical remote sensing of vegetation canopies

W. Verhoef

This report served the author as thesis to obtain the doctor's degree (PhD) at the "Landbouwniversiteit Wageningen" on January 9, 1998.

The thesis was supervised by Prof.Dr.Ir. M. Molenaar.

The content of this report may be cited on condition that full credit is given to NLR and the author.

Division:	Space
Issued:	September 1998
Classification of title:	unclassified



Abstract

Verhoef, W., 1997, Theory of radiative transfer models applied in optical remote sensing of vegetation canopies. PhD thesis, Wageningen Agricultural University, Wageningen, the Netherlands.

In this thesis the work of the author on the modelling of radiative transfer in vegetation canopies and the terrestrial atmosphere is summarized. The activities span a period of more than fifteen years of research in this field carried out at the National Aerospace Laboratory NLR.

For the interpretation of optical remote sensing observations of vegetation canopies from satellites or aircraft the use of simulation models can be an important tool, as these models give insight in the relations between vegetation properties and observed remote sensing data.

The models discussed here are first presented from a theoretical point of view. An attempt has been made to construct a framework in which all the discussed models, for vegetation as well as for the atmosphere, can be represented. After an introduction on basic radiometric quantities and relations (chapter 2), and a classification of radiative transfer models and solution methods (chapter 3), examples of existing models are discussed in chapter 4.

In chapter 5 a new generalized theory of radiative transfer models for azimuthally isotropic media is presented, namely the $(N+2)$ -stream theory. This theory describes radiative transfer in vegetation canopies or atmospheres to any desired numerical accuracy. In the formulation of the analytical solution of this model, which is based on eigenvector decomposition, much attention has been paid to possible numerical problems. Symmetry relations are exploited in order to reduce memory requirements and computation time, and expressions have been found for which the so-called reciprocity relations are automatically fulfilled. The numerical capabilities of this model for simulation of atmospheric radiative transfer are demonstrated in chapter 6.

The theory of the vegetation canopy bidirectional reflectance models SAIL and SAILH are discussed in chapters 7 and 8. Both are four-stream models. In SAILH the so-called hot spot effect, which is related with the finite leaf size, is incorporated. After these more theoretic chapters, practical aspects of the radiative transfer models are discussed in the following chapters.

In chapter 9 the application of a four-stream atmosphere model to the correction and calibration of Landsat Thematic Mapper images is



demonstrated. This atmosphere model has been coupled with the SAILH model into an overall four-stream model of optical remote sensing observations from any altitude. This model is called OSCAR (optical soil-canopy-atmosphere radiance) and is presented in chapter 10.

Applications of radiative transfer modelling to optical remote sensing problems are illustrated in chapter 11, and chapter 12 summarizes the conclusions of this thesis.

Key words: Optical remote sensing, vegetation canopies, bidirectional reflectance, radiative transfer models, radiometry, spectrometry, atmospheric correction, model inversion.



Preface

My own fascination with the subject of this thesis stems from the amazing fact that visual phenomena observed in the natural environment and in earth observation images can be simulated by means of a model in a computer. This, and the possible application of such models in the interpretation of remote sensing images, have been important drives for my work on radiative transfer models.

The opportunity to summarize my previous work on this subject, done at NLR in the period 1981-1992, has also stimulated me to complete the theory of the so-called $(N+2)$ -stream model, which I had started to investigate in 1986. During the period 1993-1995 I worked on an improved mathematical formulation of this theory and the final result, which has not been published elsewhere, is presented in this thesis. In addition, the older material presented has been thoroughly revised in order to obtain a consistent notation in which the different models can be represented.

Many people have contributed in various ways to the work presented in this thesis. First of all, I want to express my sincere thanks to my co-promoter Nico Bunnik. From the beginning of my career at NIWARS in 1973, acting as leader of the project team SpecRadVe, at NLR from 1977 to 1986 as group leader and head of the Remote Sensing Department, and from 1993 until now as my co-promoter, he encouraged me in my work on radiative transfer models. The importance of studying the fundamental "object-sensor" relationships for the interpretation and a better understanding of optical remote sensing images was always underlined by him. At the same time he kept making me aware of the need to demonstrate the practical applicability of the results of these investigations.

I am very grateful to my promoter Prof. Martien Molenaar of Wageningen Agricultural University, department of Geo-Informatics and Remote Sensing (WAU-GIRS), who suggested to me the possibility of writing a dissertation on the subject of radiative transfer models applied in optical remote sensing. I want to thank him for his willingness to act as my promoter and for the time and effort that this entails. I retain pleasant memories of the discussions we had (together with Nico Bunnik and my referee Jan Clevers) on the structure and contents of the thesis. During these sessions I got many constructive suggestions for improvement of the manuscript.

I very much appreciate the positive criticism of Jan Clevers of WAU-GIRS, and I thank him especially for his careful reading of the texts, and the



pleasant cooperation in several projects we were involved in.

Paul de Loor (retired from FEL-TNO) and Th. A. de Boer (retired from AB-DLO) were strong initiators of the NIWARS programme, and I thank them for their encouragement of the work on fundamental aspects of remote sensing. Henk van Kasteren and Dick Uenk of AB-DLO were members of the NIWARS SpecRadVe project team and I want to thank them for the pleasant cooperation in the field work and the many discussions on practical aspects of remote sensing for agricultural applications.

When Nico Bunnik and I joined the NLR in Amsterdam in 1977, Mr. L.R. Lucassen, head of the Flight Experiment and Helicopters Department, gave us a warm welcome and I enjoyed very much the pleasant discussions with him on remote sensing subjects, also because of his sincere interest. I would like to thank the successive heads of the later established Remote Sensing Department, Nico Bunnik, Allaert Kalshoven and Gert van der Burg, for their support of my continuing work on radiative transfer modelling.

I thank my former and present colleagues in the Remote Sensing Department of NLR for the pleasant cooperation through the years.

Discussions on the contents of my work took place with many people outside NLR. I thank Siegfried Gerstl of Los Alamos National Laboratory for making me aware of the work of A. Kuusk on the hot spot effect, when we met at a congress in Enschede. The discussions about atmospheric modelling with Tis Veugen, Hein van Stokkom (formerly RWS-MD), Jos Kokke (RWS-MD) and Johan de Haan (Free University Amsterdam) are much appreciated. Discussions and cooperation with Gerrit Epema of WAU and Massimo Menenti of Staring Centre - DLO on the calibration of Landsat data have contributed to an operational atmospheric correction procedure based on the four-stream atmosphere model.

The work presented in this thesis would not have been possible without the continuous support of the Board of Directors of NLR and I would like to express my sincere thanks for their willingness to keep supporting my work on radiative transfer models in the framework of NLR's own research programme, for giving me the opportunity to complete this thesis and for the support in the form of word processing and graphic art facilities.

Last but not least, I am much obliged to our division secretary Corry Scheffer for the typing of substantial parts of the manuscript and to Co Dalemans for the graphic art work on the illustrations.

Contents

Abstract	3
Preface	5
List of symbols	10
1 Introduction	17
1.1 Optical remote sensing	17
1.2 Interpretation of multispectral image data	19
1.3 Motivation and objectives	21
1.4 Organization of the thesis	23
2 Fundamental radiometric quantities	26
2.1 Radiant flux, irradiance, exitance and intensity	26
2.2 Radiance	27
2.3 Reflectance	30
2.4 Transmittance	33
2.5 Four-stream interactions with layers and surfaces	34
2.6 Extinction, absorption and scattering	37
3 Classification of radiative transfer models	44
3.1 Type of medium	44
3.2 Scattering behaviour	45
3.3 Representation of the radiative transfer equation	45
3.4 Methods of solution	47
3.5 Coupling methods	50
3.6 Conclusion	51
4 Models for vegetation and the atmosphere	52
4.1 Vegetation models	52
4.1.1 Discrete models	52
4.1.2 Analytical models	53
4.1.3 Other models	57
4.2 Atmospheric models	59
4.3 Discussion	60

5	(N+2)-Stream theory	62
5.1	Introduction	62
5.2	Tessellation	65
5.3	(N+2)-stream representation of the radiative transfer equation	72
5.4	Properties of the analytical solution	78
5.5	Mathematical validation	90
5.6	Conclusion	92
6	Simulation results for (N+2)-stream atmosphere model	94
6.1	Introduction	94
6.2	Eigenvalues and eigenvectors	94
6.3	Infinite reflectance matrix	99
6.4	Hemispherical reflectance and transmittance for specular incidence	102
6.5	Bidirectional reflectance and transmittance	111
6.6	Discussion	133
7	Light scattering by leaf layers: the SAIL model	135
7.1	Introduction	135
7.2	Extinction and scattering coefficients	138
7.3	Leaf inclination distribution functions	153
7.4	Angular profiles of bidirectional canopy reflectance	160
8	Hot spot effect in vegetation canopies	166
8.1	Introduction	166
8.2	Theory of Kuusk	166
8.3	Incorporation into the SAIL model	173
8.4	Crop growth considerations	175
8.5	Examples of simulations	176
9	Four-stream atmospheric correction model	182
9.1	Introduction	182
9.2	Description of the atmospheric effect	182
9.3	Extinction and scattering coefficients of the atmosphere	186
9.4	Model implementation aspects	192
9.5	Estimation of aerosol optical thickness	195
9.6	Atmospheric correction of Landsat Thematic Mapper images	201
9.7	Validation results	205
9.8	Conclusions	212



10 A coupled soil-vegetation-atmosphere model	213
10.1 Introduction	213
10.2 Model design	213
10.3 Model parameters	217
10.4 Implementation aspects	220
10.5 Some examples of simulation results	221
10.6 Fish-eye views	233
11 Applications	235
11.1 NOAA satellite NDVI simulation	236
11.2 Red – near infrared feature space diagrams	239
11.3 Sugar beet yield forecasting experiment	242
11.4 Model inversion	247
11.5 Estimation of fraction APAR	250
11.6 Albedo estimation	251
12 Concluding remarks	253
References	260
Summary	267
Appendix A Analytical solution (N+2)-stream radiative transfer equation	272
Appendix B Approximations for small and infinite optical thickness	307
Appendix C Extinction and scattering coefficients of the SAIL model	314

List of symbols

Symbol	Description	Unit
A	Attenuation matrix	-
A	Area	m^2
A_0, A_1	Calibration constants	$W/(m^2 \mu m sr)$
a	Attenuation coefficient	-
a	Leaf inclination distribution function parameter	-
B	Backscattering matrix	-
B_0, B_1	Calibration constants	-
b	Leaf inclination distribution function parameter	-
b	Optical thickness	-
b_A	Aerosol optical thickness	-
b_R	Rayleigh optical thickness	-
b_G	Gaseous absorption optical thickness	-
b_{H_2O}	Water vapour absorption optical thickness	-
b_{O_3}	Ozone absorption optical thickness	-
d	Layer thickness	m
d	Superscript for multiple scattering contribution	
d	Differential operator	
e	Exponential function	
E	Irradiance, radiant flux density (spectral)	$W/(m^2 \mu m)$
E_s	Solar spectral irradiance	$W/(m^2 \mu m)$
E_s^o	Solar spectral irradiance on a plane perpendicular to the sunrays	$W/(m^2 \mu m)$
E_{sun}	Solar spectral irradiance on ground level	$W/(m^2 \mu m)$
E_{sky}	Spectral irradiance from the sky on ground level	$W/(m^2 \mu m)$
E	Vector of spectral flux densities	$W/(m^2 \mu m)$
E^-, E^+	Downward and upward diffuse spectral irradiance vectors	$W/(m^2 \mu m)$
E_o^-, E_o^+	Downward and upward flux-equivalent spectral radiance	$W/(m^2 \mu m sr)$
F	Forward scattering matrix	-
$F(\theta_\ell)$	Cumulative leaf inclination distribution	-
$f(\theta_\ell)$	Leaf inclination density	rad^{-1}
f_s, f_o	Leaf area projection factors in direction of sun and observation	-

Symbol	Description	Unit
f_1, f_2	Leaf area projection factors for diffuse fluxes	-
H	Atmospheric height	km
H_0, H_1, H_2	Atmospheric scale heights	km
h	Height	m, km
I	Identity matrix	
I	Spectral radiant intensity	$W/(\mu m \text{ sr})$
i	Imaginary unit	
i	Incidence zenith zone number	
j	Exitence zenith zone number	
K	Extinction coefficient in the direction of observation	-
k	Extinction coefficient in the direction of the sun	-
k	Azimuth sector number	
L	Leaf area index (LAI)	-
L'	Leaf area density	m^{-1}
L	Spectral radiance	$W/(m^2 \mu m \text{ sr})$
L_i	Incident spectral radiance	$W/(m^2 \mu m \text{ sr})$
L_o	Spectral radiance in direction of observation	$W/(m^2 \mu m \text{ sr})$
$\mathcal{L}$	Spectral leaf radiance	$W/(m^2 \mu m \text{ sr})$
ℓ	Horizontal correlation length	m
ℓ	Leaf normal vector	
M	Coefficient matrix	-
M	Spectral exitance	$W/(m^2 \mu m)$
m	Azimuthal mode of discrete cosine transform	
N	Number of diffuse flux angular segments	
N_Z	Number of zenith zones	
N_A	Number of azimuth sectors	
O	Zero matrix	
o	Vector in direction of observation	
P	Cumulative probability	
P_s	Probability of solar illumination	
P_o	Probability of direct line of sight	
P_{so}	Joint probability of viewing sunlit canopy elements	
p	Scattering phase function	sr^{-1}
R	Diffuse reflectance matrix	-
r_o	Directional reflectance factor	-



Symbol	Description	Unit
r_d	Diffuse (hemispherical) reflectance factor	-
r_p	Planetary reflectance	-
r_t	Target reflectance	-
r_s	Soil reflectance	-
r_b	Background reflectance	-
r_{so}	Bidirectional reflectance	-
r_{sd}	Diffuse reflectance for specular incidence	-
r_{do}	Directional reflectance for diffuse incidence	-
r_{dd}	Diffuse (hemispherical) reflectance for diffuse incidence	-
s	Diffuse backscattering vector for specular incidence	-
s'	Diffuse forward scattering vector for specular incidence	-
s	Superscript for single scattering contribution	
s	Diffuse backscattering coefficient for specular incidence	-
s'	Diffuse forward scattering coefficient for specular incidence	-
T	Diffuse transmittance matrix	-
T	Transposed (superscript)	
$T_1 T_2$	Two-way total transmittance	-
V	Visibility	km
ν	Directional backscattering vector for diffuse incidence	-
ν'	Directional forward scattering vector for diffuse incidence	-
ν	Directional backscattering coefficient for diffuse incidence	-
ν'	Directional forward scattering coefficient for diffuse incidence	-
W	Discrete cosine transform matrix	
w	Bidirectional scattering coefficient for upward existence direction	-
w'	Bidirectional scattering coefficient for downward existence direction	-

Symbol	Description	Unit
x	Relative optical height	-
Y	Eigenvector matrix	-
Z	Azimuthal scattering matrix	-
z	Metrical depth	m
α	Geometric factor in SAILH model	-
α	Ångström coefficient	-
β	Turbidity factor	-
β	Interception coefficient	m^{-1}
β_A	Aerosol interception coefficient	m^{-1}
β_R	Rayleigh interception coefficient	m^{-1}
γ	Bidirectional scattering coefficient	m^{-1}
γ'	Volume scattering function	m^{-1}
Δ	Difference operator	
δ	Angular difference	rad, °
δ	Scattering angle	rad, °
δ	Dirac delta function	
δ_1, δ_2	Boundary equation vectors	$W/(m^2 \mu m)$
ϵ	Subscript for small optical thickness	
∞	Subscript for infinite optical thickness	
η_A	Aerosol backscattering efficiency	-
θ	Zenith angle	rad, °
θ_i	Incidence zenith angle	rad, °
θ_o	Observation zenith angle	rad, °
θ_s	Solar zenith angle	rad, °
κ	Diffuse extinction matrix (diagonal)	-
κ	Diffuse extinction coefficient	-
Λ	Eigenvalue matrix (diagonal)	-
λ	Eigenvalue	-
λ	Wavelength	$\mu m, nm$
μ	$\cos \theta$	
ν	Normalized spectral radiance	-
ρ	Horizontal spatial correlation coefficient	-
ρ_L	Single leaf reflectance	-
ρ_s	Soil reflectance	-
ρ	Lambertian reflectance	-
ρ'	Bidirectional reflectance distribution function	sr^{-1}

Symbol	Description	Unit
ρ_{so}	Bidirectional reflectance of a layer	-
ρ_{sd}	Diffuse reflectance for specular incidence of a layer	-
ρ_{do}	Directional reflectance for diffuse incidence of a layer	-
ρ_{dd}	Diffuse reflectance for diffuse incidence of a layer	-
ρ_{sd}	Diffuse reflectance vector for specular incidence of a layer	-
ρ_{do}	Directional reflectance vector for diffuse incidence of a layer	-
σ	Diffuse backscattering coefficient for diffuse incidence	-
σ'	Diffuse forward scattering coefficient for diffuse incidence	-
τ_L	Single leaf transmittance	-
τ	Lambertian transmittance	-
τ	Optical depth	-
τ'	Bidirectional transmittance distribution function	sr ⁻¹
τ_{so}	Bidirectional transmittance	-
τ_{sd}	Diffuse transmittance for specular incidence	-
τ_{do}	Directional transmittance for diffuse incidence	-
τ_{dd}	Diffuse transmittance for diffuse incidence	-
τ_{ss}	Direct transmittance in direction of the sun	-
τ_{oo}	Direct transmittance in direction of observation	-
τ_{ssoo}	Two-way direct transmittance	-
τ_{sd}	Diffuse transmittance vector for specular incidence	-
τ_{do}	Directional transmittance vector for diffuse incidence	-
Φ	Radiant power flux	W
Φ_λ	Spectral radiant power flux	W/ μ m
ψ	Absolute azimuth difference	rad, °
Ω	Solid angle	sr
Ω_i	Incidence solid angle	sr
Ω_o	Exitence solid angle	sr
ω	Single scattering albedo	-
ω_A	Aerosol single scattering albedo	-



Symbol	Description	Unit
φ	Azimuth angle	rad, °
φ_i	Incidence azimuth angle	rad, °
φ_o	Observation azimuth angle	rad, °
φ_s	Solar azimuth angle	rad, °



This page is intentionally left blank



1 INTRODUCTION

The subject of this thesis is the theory of radiative transfer models that can be applied to the simulation of the observation of vegetation canopies by means of optical remote sensing techniques.

In this introduction first a brief review of optical remote sensing techniques is given. Next, the physical interpretation of the spectroradiometric content of digital earth observation imagery is discussed. Radiative transfer models can be applied to assist in this interpretation. The motivation and objectives for writing this thesis are summarized next and this chapter is concluded with a structural overview of the remaining chapters.

1.1 Optical remote sensing

Optical remote sensing refers to the detection of electromagnetic radiation by means of electronic sensors or photographic material in the part of the spectrum where the main phenomena are similar to those in the visible region.

In this spectral region the sun is the primary source of incoming radiation and the radiation detected by a remote sensing device consists mainly of radiation reflected at the earth's surface or scattered by the atmosphere. When the sun is used as the radiation source one speaks of a *passive* optical remote sensing technique. An example of an *active* optical remote sensing technique is LIDAR (LIght Detection And Ranging). Here the source (usually a laser) is carried on board an aircraft and light emission and detection of the reflected or scattered radiation both take place in the aircraft. In the remainder of this thesis only passive optical remote sensing is considered.

The part of the spectrum which is commonly referred to as the optical region ranges from wavelengths of about 400 nm to about 2500 nm ($1 \text{ nm} = 10^{-9} \text{ m}$). Shorter wavelengths are rarely used because of the rapidly decreasing intensity of the sun in the ultraviolet and the very strong scattering of radiation in the atmosphere, which makes it difficult to observe the earth's surface. At longer wavelengths than $2.5 \mu\text{m}$ ($1 \mu\text{m} = 10^{-6} \text{ m}$) the solar radiation also becomes weaker and major water vapour absorption bands prevent the solar radiation from reaching the ground. Besides, thermal emission of earth and atmosphere become increasingly important towards



longer wavelengths, and are the dominant sources of radiation in the atmospheric windows of 3-5 μm and 8-14 μm . The latter are usually referred to by the name thermal infrared.

The optical spectral region is usually divided into a number of subregions as follows:

Wavelength region	Name
400 - 700 nm	Visible
700 - 1300 nm	Near infrared
1300 - 2500 nm	Shortwave infrared

Sometimes the latter region is called mid infrared, but there is no general consensus on this, so the term shortwave infrared is preferred, although the wavelengths of the near infrared are shorter.

The visible region is the one to which the human eye is sensitive. The retina of the human eye contains biological detectors sensitive to blue, green and red light, thus making colour vision possible.

Near infrared radiation was first imaged shortly after World War II by means of photographic film which was sensitive up to about 900 nm. This type of "false colour" film makes it possible to record the strong reflection of vegetation canopies and trees in the near infrared, which is expressed by an intense red colour.

Imaging of electromagnetic radiation at wavelengths beyond 900 nm became only possible with the advent of electronic detectors made of semiconducting materials such as silicon, germanium, lead sulphide, etcetera. In the sixties this has resulted in the development of opto-mechanical multispectral scanners operated from an aircraft in order to image the earth's surface. These instruments make use of a rotating or oscillating mirror for scanning. The collected radiation is first decomposed into its spectral constituents by means of dispersion elements like beam splitters, prisms and gratings, and subsequently focused onto several detectors. The across-track scanning provided by the mirror and the forward motion of the platform together form the means to build up an image in several spectral bands simultaneously. The electric signals from the detectors are digitized and recorded on high density magnetic tape. Digital recording of multispectral images has several advantages, such as reliable archiving and the enormous possibilities for process-



ing the image data by means of a computer.

With the exploration of the moon and planets other than the earth by means of satellites one had no alternative but to use electronic sensing and transmitting the digital images to receiving stations on earth. The development of this technique has also led to the launch of the first remote sensing satellite (Landsat 1, in 1972) which was intended to map the earth's resources and the environment. Before this, weather satellites using a similar data acquisition system had already been launched.

The first SPOT satellite, launched in 1986, contains a so-called pushbroom CCD-array scanner, called HRV instrument. In this case, there are no moving parts associated with scanning, but rather across-track scanning takes place fully electronically by sequentially reading out the elements of the linear CCD detector arrays. This gives a high geometric image quality. The HRV instrument is also across-track pointable, which gives more flexibility in the selection of targets and allows to obtain stereo pairs of images.

At present, so-called hyperspectral sensors or imaging spectrometers using detector arrays have been tested on board aircraft and plans exist to launch such instruments on board earth observation satellites in the near future. These instruments are capable of recording the radiation from the earth's surface in hundreds of spectral bands.

1.2 Interpretation of multispectral image data

The development of advanced multispectral remote sensing systems for earth observation such as reviewed in the previous section and the study of the enormous amounts of resulting image material both have stimulated the research in the field of image interpretation of this kind of imagery.

The interpretation of the radiometric content of digital multispectral data concerns the question what information is contained in the digital numbers representing the signals detected by the sensor in the various spectral bands. All the sensing devices mentioned in section 1.1, from the human eye and photographic film to the most advanced hyperspectral sensors, respond to a physical radiometric quantity called spectral radiance, with the symbol L . The term radiance has been derived from an equivalent quantity in visual perception theory called luminance or brightness. The human eye and photographic film respond directly to spectral radiance inputs, but in a non-linear fashion. However, after calibration, electronic sensors allow to obtain the



quantitative spectral radiance value.

The spectral radiance detected by an earth observation instrument on board an aircraft or a satellite is not equal to the spectral radiance of the ground area element (pixel) on the earth's surface which is inside the instantaneous field-of-view of the sensor. This would only be the case if there were no atmosphere. The atmosphere between target and sensor modifies the radiance because the radiance coming from the ground is attenuated on its path through the atmosphere and scattered atmospheric radiation is added. The simplest equation (also called the remote sensing equation) to describe this effect is given by

$$L_s = L_p + TL_t \quad , \quad (1.1)$$

where L_t is the target's (spectral) radiance in the direction of the sensor, T is the direct transmittance from target to sensor, L_p is the so-called path radiance and L_s is the radiance sensed by the instrument.

Information on the characteristics of the target is contained in the target's radiance, L_t . However, this quantity depends also on the total downward radiant flux density incident on the target, called total irradiance, E_{tot} . If the target can be considered as a so-called Lambertian reflecting surface (a perfect diffusor), then the target's radiance L_t is equal in all directions and given by

$$L_t = \pi^{-1} \rho E_{tot} \quad , \quad (1.2)$$

where ρ is the Lambertian reflectance of the target. The factor π^{-1} is not essential, but merely follows from the physical units in which the irradiance and the radiance are defined. In chapter 2 this is explained in more detail.

In the above simple model of optical remote sensing, given by Eqs. (1.1) and (1.2), the surface reflectance ρ is the quantity that depends only on the properties of the target, so in general this will be the quantity of interest. From reflectance values in several spectral bands one should be able to draw conclusions on the properties of the target. However, for this the relations between an object's properties and the resulting reflectance spectrum are needed. Such a set of relations is called a reflectance model. By means of a reflectance model one can predict spectral reflectances from a number of object parameters and by means of model inversion techniques one can derive one or more object parameters from given spectral reflectance data.



Most objects in reality are non-Lambertian, which means that their reflectance is direction-dependent. In general the reflectance of an object depends on the direction of the incoming radiation as well as on the direction of the reflected radiation, or, in other words, the reflectance is a bidirectional quantity. Besides that this considerably complicates the description of optical remote sensing given above, it also offers the opportunity to make use of it, for instance by combining observations under different viewing angles, which gives additional information about the object's properties. In that case a bidirectional reflectance model can assist in the interpretation of spectral-directional reflectance data.

Reflectance models are mostly based on radiative transfer theory, i.e. the theory that describes all interactions of radiation with a medium by means of mathematical-physical relations. In these theories the medium usually is considered to consist of a plane-parallel layer of absorbing and scattering material. This is also the case for all models discussed in this thesis.

The observation of vegetation canopies by means of optical remote sensing can be modelled with a bidirectional vegetation reflectance model in combination with an atmospheric radiative transfer model. Such a combination of models can be very useful as it allows to lay a direct relation between vegetation properties and observed remote sensing data.

In this thesis differences and similarities between radiative transfer models for vegetation canopies and atmospheres will be explored in order to establish a framework in which both types of model can be expressed. Using this framework, and applying the so-called adding method (Van de Hulst, 1980), it is possible to integrate both models into a single model of the observation of vegetation canopies by optical remote sensing.

1.3 Motivation and objectives

The author's involvement in radiative transfer modelling began in 1974 when, in the framework of the NIWARS research programme (cf. Bunnik, 1978), large amounts of reflectance measurement data from agricultural crops were acquired by means of a field spectrometer. One of the means to interpret the measured reflectance spectra was the use of a vegetation canopy reflectance model for simulating the measurements. For this the Suits model (Suits, 1972) was applied, because it was easy to use, relatively simple and nevertheless could calculate bidirectional spectral reflectances. Besides this



application, the model was also extensively used for sensitivity analyses and the study of vegetation indices (Bunnik, 1978). An extension of the Suits model in order to accommodate also crops with a clear row structure was proposed by the author in 1975 (Verhoef & Bunnik, 1976).

At the National Aerospace Laboratory NLR, since 1977, the author has continued the work on radiative transfer modelling. The SAIL model (Verhoef & Bunnik, 1981; Verhoef, 1984) was developed as a refinement of Suits's model in order to obtain more realistic bidirectional reflectance profiles as a function of the viewing angle. The differential equations on which the SAIL and Suits models are based could be identified as more general radiative transfer equations, part of a so-called four-stream theory (Verhoef, 1985a). In the four-stream framework also the radiative transfer in the atmosphere could be expressed quite easily, and the coupling of vegetation bidirectional reflectance with atmospheric models could be realized by means of the adding method of Van de Hulst. These developments have resulted in a so-called scene radiation model (Verhoef, 1985b) which could be used for the simulation of remote sensing observations of vegetation canopies from space.

Four-stream radiative transfer theory is based on the existence of a direct solar flux, two hemispherical (semi-isotropic) diffuse fluxes (up- and downward) and a flux (radiance) in the direction of viewing. A disadvantage of this concept is that in principle the calculation of the contribution due to multiple scattering is inaccurate, since with each scattering event any directional flux is immediately converted into two semi-isotropical fluxes, and this is not realistic, certainly not for atmospheres, in which aerosol scattering is strongly peaked in the forward direction. In order to refine the four-stream concept in this respect, a so-called $(N+2)$ -stream theory has been developed by the author. This theory is based on the same two directional fluxes as before (in the solar and viewing directions), but the two hemispherical fluxes are replaced by N diffuse fluxes which are only isotropic within certain angular segments on the sphere of directions.

Since the presentation of the first results of the $(N+2)$ -stream model for atmospheric scattering (Verhoef, 1988) the mathematical formulation of the analytical solution of this model has been revised several times, mainly because the first formulations resulted into numerically unreliable computer code. These difficulties have been solved now and the results are presented in chapters 5 and 6 and Appendices A and B of this thesis.

An improvement of four-stream theory in a completely different direction is



the incorporation of the hot spot effect in the SAIL model. This effect is caused by the finite size of the leaves and cannot be described in a traditional radiative transfer model. As a matter of fact, the (traditional) radiative transfer equation does not contain any parameters that depend on the size of the scattering elements. What is needed in this case is a close reexamination of the process of single scattering. Such an analysis has been performed by Kuusk (1985). His theory considers the statistical dependence between the probabilities of direct illumination by the sun and direct line-of-sight from outside the canopy. The joint probability of these two events forms the basis of the single scattering contribution to the bidirectional reflectance of a leaf canopy.

An important motivation for writing this thesis was that a unified theoretic framework, in which various radiative transfer models can be presented, has been completed now. Also, the author is convinced that, certainly with the advent of more complex earth observation instruments like imaging spectroradiometers, for the quantitative interpretation of optical remote sensing data radiative transfer models will prove to be an indispensable tool.

The objectives can be summarized as follows:

- Explication of radiative transfer theory based on classical radiometric relations
- Formulation of a unified theoretic concept applicable to radiative transfer in atmospheres as well as vegetation canopies
- Classification of radiative transfer models on the basis of mathematical and physical features within the framework of the unified theory
- Introduction to the theory of the (N+2)-stream model and demonstration of its results
- Demonstration of applications of radiative transfer modelling in optical remote sensing

1.4 Organization of the thesis

In chapter 2 the radiometric quantities and relationships forming the basis of radiative transfer models are introduced. Important physical quantities are (spectral) irradiance, radiance and various reflectance and transmittance



quantities. Extinction and scattering coefficients are the prime quantities to describe radiative transfer in plane-parallel layers. Optical thickness, the single scattering albedo and the scattering phase function are quantities which are mostly used in the description of radiative transfer in the atmosphere. Many of the quantities and relationships introduced here are applied in later chapters.

A classification of radiative transfer models in terms of a number of physical features is proposed in chapter 3. Based on these features, a number of existing radiative transfer models for vegetation canopies and atmospheres are briefly discussed in chapter 4.

Chapter 5 and Appendices A and B together provide a complete theoretical description of the $(N+2)$ -stream model, which is a generalization of the four-stream model. In the mathematical formulation of the $(N+2)$ -stream model several aspects are discussed, such as the use of symmetry relations to simplify the analytical solution, an azimuthal cosine transform and numerical problems related with equal eigenvalues. The mathematical validity of the model is verified by means of the so-called reciprocity relations and by approximations for small and infinite optical thickness. In chapter 6 a large number of examples of simulations carried out with this model are presented in order to demonstrate its performance under a wide variety of input conditions.

The $(N+2)$ -stream concept as presented in chapter 5 is applicable to any azimuthally isotropic medium, thereby including atmospheres as well as vegetation canopies with uniform leaf azimuth distribution and random leaf placement. The four-stream SAIL model is an example of such a model for leaf canopies that is based on these same assumptions. It is presented in chapter 7 and Appendix C, with emphasis on the derivation of the extinction and scattering coefficients and their dependence on the leaf angle distribution.

In chapter 8 the hot spot effect in the bidirectional reflectance of leaf canopies is discussed. As mentioned already in the previous section, this is caused by the finite size of the leaves and here a physical description of the effect is given, following the theory of Kuusk (1985). This theory has been incorporated in the SAIL model in 1989, and the resulting model is called SAILH. The SAIL model now forms a special case of SAILH, which can be obtained by setting the leaf size equal to zero. Some implications of the hot spot effect in relation with canopy morphology and crop growth are discussed on the basis of simulation results.



After these more theoretic previous chapters, practical aspects of the radiative transfer models are discussed in the following chapters.

In chapter 9 the application of a four-stream atmosphere model to the correction and calibration of Landsat Thematic Mapper images is demonstrated. This atmosphere model has been coupled with the SAILH model into an overall four-stream model of optical remote sensing observations from any altitude. This model is called OSCAR (optical soil-canopy-atmosphere radiance) and is presented in chapter 10. The so-called adjacency effect is also included in this model.

A number of applications of radiative transfer modelling to optical remote sensing problems are illustrated in chapter 11. The applications are divided into the categories prediction, explication and model inversion.

Chapter 12 summarizes the conclusions of this thesis and recommendations for further work are given.

2 FUNDAMENTAL RADIOMETRIC QUANTITIES

This chapter introduces a number of radiometric quantities and definitions which will be referred to in most of the subsequent chapters. Here the term radiometric is meant to refer to electromagnetic radiation. Part of the material discussed is based on the excellent review of this subject in the "Manual of Remote Sensing", chapter 2, called "The Nature of Electromagnetic Radiation" (Colwell, 1983).

By analogy with the words incident and incidence, which refer to incoming radiation, the words "exitent" and "exitence" are introduced here in order to indicate radiation leaving an object. In this way one can speak of exitent flux and a direction of exitence, independent of the cause (reflection, transmission or emission).

The term "exitance", which is also introduced in this chapter, is not new and refers to the flux density of radiation leaving a surface, and is a more neutral alternative for the formerly used term "emittance", which has a strong association with thermal emission.

2.1 Radiant flux, irradiance, exitance and intensity

The term radiant flux, Φ , is defined as "the time rate with which radiant energy passes a spatial position" (Colwell, 1983). It is given in units of joules/sec or watts (W).

From the need to distinguish radiation according to its wavelength, its direction and the area it is incident on or reflected from, several other quantities have emerged, which are all based on radiant flux. The term spectral flux Φ_λ , for instance, is defined as the radiant flux per unit wavelength interval, with the defining equation

$$\Phi_\lambda = d\Phi / d\lambda \quad , \quad (2.1)$$

and its unit is W / μm or W / nm.

In the remainder of this thesis, it is always assumed that the radiometric quantities are spectral, unless stated otherwise, and the subscript λ will be omitted. However, units will be given in their spectral form and the unit of wavelength interval used will be the μm (10^{-6} m). In addition, the



polarization of the electromagnetic radiation will be ignored. Although it can be an extra source of information, it makes the theory considerably more complex, and its use for earth observation is still in the experimental stage.

Radiant flux density is the radiant flux per unit area. In order to distinguish between flux incident on a surface and flux leaving a surface two symbols are used, namely E for incident flux density, also called irradiance, and M for emitted or reflected flux density, also called exitance. In both cases the unit is $W / (m^2 \mu m)$. The defining equations are

$$E = d\Phi_i / dA \quad \text{and} \quad M = d\Phi_e / dA \quad , \quad (2.2)$$

where Φ_i and Φ_e denote incident and exitent flux, respectively, and dA is an infinitesimally small area.

Radiant flux intensity, I , is the radiant flux per unit solid angle. This term is often used when it is necessary to distinguish flux according to its direction. The defining equation is

$$I = d\Phi / d\Omega \quad , \quad (2.3)$$

where $d\Omega$ denotes an infinitesimal interval of solid angle. The unit of (spectral) intensity is $W / (\mu m \text{ sr})$. In the literature on remote sensing the term intensity is not frequently used, perhaps because it applies more to point sources than to extended sources, for which the term radiance seems more appropriate.

2.2 Radiance

Radiance, L , is defined as "the radiant flux per unit solid angle leaving an extended source in a given direction per unit projected source area in that direction" (Colwell, 1983). The concept of radiance is intended to correspond to the term brightness in visual perception. Its defining equation is

$$L = d^2\Phi / (d\Omega dA \cos \delta) \quad , \quad (2.4)$$

where δ is the angle between the surface's normal and the direction of flux.



The unit of (spectral) radiance is $W / (m^2 \mu m sr)$.

The definition in words given above is somewhat more strict than necessary, since the concept of radiance can equally well be used for the description of flux intercepted from, for instance, a part of the sky. In that case dA is the area intercepting the radiation and $d\Omega$ is the solid angle of the part of the sky as seen from the surface. This explains why one can speak of the radiance of the sky (or a part thereof), while this is a source that is extended but has no definite area. Radiance is an important quantity because the signal measured by electronic sensors, the exposure of photographic material, or the response of the human eye, are all a direct function of the object's radiance.

Combining the definitions of radiance and flux density, one may write

$$L_i = dE / (\cos \delta d\Omega) \text{ or } dE = L_i \cos \delta d\Omega, \text{ and}$$

$$L_o = dM / (\cos \delta d\Omega) \text{ or } dM = L_o \cos \delta d\Omega,$$

where the former equations describe the irradiance contribution from a fraction within a solid angle $d\Omega$ of incident radiance L_i and the latter describe how the radiance L_o from a surface in a certain direction contributes to the total exitance of that surface.

For the indication of directions use is made of a spherical co-ordinate system with the polar angles θ (the zenith angle) and φ (the azimuth angle) as shown in Fig. 2.1.

For a horizontal surface element placed in the centre of the sphere, $\delta = \theta$, and $d\Omega = \sin \theta d\theta d\varphi$. In this case one can calculate the irradiance on the surface element if the radiance distribution over the upper hemisphere, $L(\theta, \varphi)$, is given:

$$dE = L(\theta, \varphi) \cos \delta d\Omega = L(\theta, \varphi) \cos \theta \sin \theta d\theta d\varphi, \text{ or}$$

$$E = \int_0^{2\pi} \int_0^{\pi/2} L(\theta, \varphi) \cos \theta \sin \theta d\theta d\varphi.$$

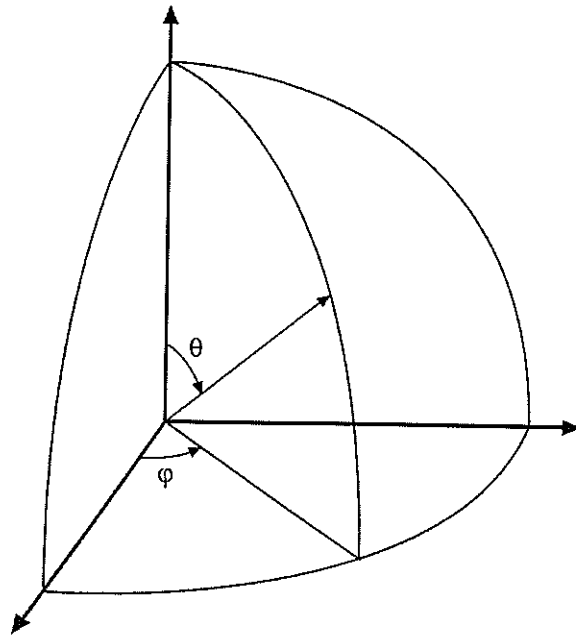


Fig. 2.1 Spherical co-ordinate system for indication of directions

In the literature $|\cos \theta|$ is usually replaced with the symbol μ , so that for $\theta < \pi/2$ $\sin \theta d\theta = -d \cos \theta = -d\mu$, and

$$E = \int_0^{2\pi} \int_1^0 -L(\mu, \varphi) \mu d\mu d\varphi = \int_0^{2\pi} \int_0^1 L(\mu, \varphi) \mu d\mu d\varphi.$$

If $L(\mu, \varphi)$ is constant over the upper hemisphere, then

$$E = L 2\pi \left. \frac{1}{2} \mu^2 \right|_0^1 = \pi L.$$

Similarly, one obtains $M = \pi L$ for the exitance of a surface having a constant radiance L (independent of direction).

For a parallel beam of (specular) incident flux $d\Omega$ is zero, so in that case the term radiance is of no use, since it would be infinite. However, the incident flux density remains finite and will be denoted as E_s in order to indicate that the irradiance is specular. The direction of a specular source of radiant flux will be indicated as (μ_s, φ_s) .



2.3 Reflectance

In general, reflectance is the ratio of reflected and incident radiant flux. The reflectance of a surface depends on the direction of observation (μ_o, φ_o) and on the angular distribution of the incident radiation. The basic quantity describing the reflectance of a surface is the bidirectional reflectance distribution function (BRDF), for which the symbol ρ' is used here. The BRDF relates reflected radiance to incident radiance by (Nicodemus, 1970 and Kasten & Raschke, 1974)

$$dL_o(\mu_o, \varphi_o) = \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\Omega_i, \quad (2.5)$$

and for specular incident flux from direction (μ_s, φ_s) it is expressed as

$$L_o(\mu_o, \varphi_o) = \rho'(\mu_s, \varphi_s, \mu_o, \varphi_o) E_s(\mu_s, \varphi_s). \quad (2.6)$$

For a Lambertian reflector ρ' is a constant, which means that in that case the radiance L_o of the surface will also be constant (independent of the direction (μ_o, φ_o)). The exitance M of a Lambertian reflector is given by

$$M = \int_0^{2\pi} \int_0^1 L_o \mu_o d\mu_o d\varphi_o = \pi L_o,$$

and for a so-called white Lambertian reflector the exitance M_{WL} equals the total irradiance E_i , since in that case all incident flux is reflected. Its radiance, L_{WL} , is given by $L_{WL} = M_{WL} / \pi = E_i / \pi$, so that the BRDF of a white Lambertian surface is equal to $\rho'_{WL} = \pi^{-1}$.

The radiance $L_o(\mu_o, \varphi_o)$ of a surface with arbitrary BRDF is found by integrating Eq. (2.5) over all incidence angles:

$$L_o(\mu_o, \varphi_o) = \int_0^{2\pi} \int_0^1 \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i. \quad (2.7)$$

The ratio of this radiance and the radiance of a white Lambertian surface with identical orientation and under identical incident radiation is called directional reflectance, with the symbol r_o . Since

$$L_{WL} = E_i / \pi = \pi^{-1} \int_0^{2\pi} \int_0^1 L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i \quad , \quad r_o \text{ is given by}$$

$$r_o = \frac{L_o(\mu_o, \varphi_o)}{L_{WL}} = \frac{\pi \int_0^{2\pi} \int_0^1 \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i}{\int_0^{2\pi} \int_0^1 L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i} \quad . \quad (2.8)$$

For specular incident flux this reduces to the so-called bidirectional reflectance, r_{so} , given by

$$r_{so} = \pi \rho'(\mu_s, \varphi_s, \mu_o, \varphi_o) = \pi L_o(\mu_o, \varphi_o) / E_s \quad . \quad (2.9)$$

For isotropic diffuse incident flux (L_i constant) one defines the directional reflectance for diffuse incidence, r_{do} , again as a ratio of radiances $L_o(\mu_o, \varphi_o) / L_{WL}$, by

$$r_{do}(\mu_o, \varphi_o) = \int_0^{2\pi} \int_0^1 \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) \mu_i d\mu_i d\varphi_i = \pi L_o(\mu_o, \varphi_o) / E_i \quad . \quad (2.10)$$

This reflectance is related to r_{so} by

$$r_{do}(\mu_o, \varphi_o) = \pi^{-1} \int_0^{2\pi} \int_0^1 r_{so}(\mu_s, \varphi_s, \mu_o, \varphi_o) \mu_s d\mu_s d\varphi_s \quad . \quad (2.11)$$

The ratio of the exitance M of a surface and the exitance M_{WL} of a white Lambertian surface with identical orientation and under identical incident radiation is called diffuse reflectance, with the symbol r_d . The exitance M of a surface with arbitrary BRDF is found by integrating its radiance $L_o(\mu_o, \varphi_o)$ over all angles of observation:

$$M = \int_0^{2\pi} \int_0^1 L_o(\mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o \quad .$$

Substitution of Eq. (2.7) and division by M_{WL} gives for r_d :

$$r_d = \frac{\int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i \mu_o d\mu_o d\varphi_o}{\int_0^{2\pi} \int_0^1 L_i(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i} \quad . \quad (2.12)$$

For specular incident flux this reduces to the so-called diffuse reflectance for specular incidence, r_{sd} , which is given by

$$r_{sd}(\mu_s, \varphi_s) = \int_0^{2\pi} \int_0^1 \rho'(\mu_s, \varphi_s, \mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o \quad . \quad (2.13)$$

This reflectance is related to r_{so} by

$$r_{sd}(\mu_s, \varphi_s) = \pi^{-1} \int_0^{2\pi} \int_0^1 r_{so}(\mu_s, \varphi_s, \mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o \quad . \quad (2.14)$$

For isotropic diffuse incident flux one finds the diffuse reflectance for diffuse incidence, r_{dd} , as the ratio M/M_{WL} , by

$$r_{dd} = \pi^{-1} \int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) \mu_i d\mu_i d\varphi_i \mu_o d\mu_o d\varphi_o \quad . \quad (2.15)$$

This reflectance can be related to r_{do} and r_{sd} by using Eqs. (2.10) and (2.13), which gives



$$r_{dd} = \pi^{-1} \int_0^{2\pi} \int_0^1 r_{do}(\mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o = \pi^{-1} \int_0^{2\pi} \int_0^1 r_{sd}(\mu_s, \varphi_s) \mu_s d\mu_s d\varphi_s \quad . \quad (2.16)$$

The interaction of incident radiation with a horizontal surface can be considerably simplified if one assumes that the incident flux is composed of a specular part E_s and a diffuse part E^- which is isotropic. In that case the radiance L_o of the surface in a direction (μ_o, φ_o) is found from

$$\pi L_o = r_{so} E_s + r_{do} E^- \quad . \quad (2.17)$$

For the exitance M of the surface, here denoted as an upward diffuse flux density E^+ , this approximation gives

$$E^+ = r_{sd} E_s + r_{dd} E^- \quad . \quad (2.18)$$

Eqs. (2.17) and (2.18) constitute the four-stream representation of the reflectance of a surface, as four streams or fluxes are involved. In remote sensing this is useful, since to a first approximation and under clear conditions the radiation from the sky is isotropic and the sun is a source of specular flux. The term πL_o can be related to the signal measured by remote sensing instruments, whereas the term E^+ is relevant for the interaction between the atmosphere and the surface, and for e.g. energy balance studies.

2.4 Transmittance

The transmittance of radiant flux through a plane parallel layer can be described in close analogy with that given for the reflection at a surface in the previous section. All diffuse transmittances so defined can be based on a so-called bidirectional transmittance distribution function BTDF, for which the symbol τ' is used. The BTDF relates transmitted radiance to incident radiance by

$$dL_o(\mu_o, \varphi_o) = \tau'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\Omega_i \quad . \quad (2.19)$$

The difference with Eq. (2.5) is that in Eq. (2.5) the directions (μ_i, φ_i) and (μ_o, φ_o) are in opposite hemispheres, whereas in Eq. (2.19) these must be in



the same hemisphere.

In a manner similar to the one applied for the reflectances defined in section 2.2, the following transmittances can be defined:

$$\tau_{so}(\mu_s, \varphi_s, \mu_o, \varphi_o) = \pi \tau'(\mu_s, \varphi_s, \mu_o, \varphi_o) \quad ; \quad (2.20)$$

$$\tau_{do}(\mu_o, \varphi_o) = \pi^{-1} \int_0^{2\pi} \int_0^1 \tau_{so}(\mu_s, \varphi_s, \mu_o, \varphi_o) \mu_s d\mu_s d\varphi_s \quad ; \quad (2.21)$$

$$\tau_{sd}(\mu_s, \varphi_s) = \pi^{-1} \int_0^{2\pi} \int_0^1 \tau_{so}(\mu_s, \varphi_s, \mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o \quad ; \quad (2.22)$$

$$\tau_{dd} = \pi^{-1} \int_0^{2\pi} \int_0^1 \tau_{do}(\mu_o, \varphi_o) \mu_o d\mu_o d\varphi_o = \pi^{-1} \int_0^{2\pi} \int_0^1 \tau_{sd}(\mu_s, \varphi_s) \mu_s d\mu_s d\varphi_s \quad . \quad (2.23)$$

For $\mu_o = \mu_s$ and $\varphi_o = \varphi_s$, radiation can penetrate into a layer and leave it at the opposite side without any interaction. The fraction of the incident flux which passes through the layer in this manner is called the direct transmittance. For specular flux it is symbolized as τ_{ss} , whereas for radiance in the direction of observation it is symbolized as τ_{oo} .

2.5 Four-stream interactions with layers and surfaces

The quantities describing four-stream radiant flux interactions with a surface were given in section 2.3. For a plane-parallel layer the six transmittances defined in the previous section are used in addition to four reflectances which describe the reflections at the top and the bottom of the layer and which are symbolized as ρ , with the proper subscripts attached. In that case the four-stream radiative transfer equations for a layer are the following:



$$E_s(b) = \tau_{ss} E_s(t) \quad , \quad (2.24.a)$$

$$E^-(b) = \tau_{sd} E_s(t) + \tau_{dd} E^-(t) + \rho_{dd} E^+(b) \quad , \quad (2.24.b)$$

$$E^+(t) = \rho_{sd} E_s(t) + \rho_{dd} E^-(t) + \tau_{dd} E^+(b) \quad , \quad (2.24.c)$$

$$\pi L_o^+(t) = \rho_{so} E_s(t) + \rho_{do} E^-(t) + \tau_{do} E^+(b) + \tau_{oo} \pi L_o^+(b) \quad , \quad (2.24.d)$$

$$\pi L_o^-(b) = \tau_{so} E_s(t) + \tau_{do} E^-(t) + \rho_{do} E^+(b) + \tau_{oo} \pi L_o^-(b) \quad , \quad (2.24.e)$$

where (t) and (b) refer to top and bottom of the layer, and

E_s = specular downward irradiance

E^- = diffuse downward irradiance

E^+ = diffuse upward irradiance

L_o^+ = upward radiance

L_o^- = downward radiance

τ_{ss} = direct transmittance in direction of specular flux

τ_{oo} = direct transmittance in direction of observed radiance

τ_{so} = bidirectional transmittance

τ_{sd} = diffuse transmittance for specular incidence

τ_{do} = directional transmittance for diffuse incidence

τ_{dd} = diffuse transmittance for diffuse incidence

ρ_{so} = bidirectional reflectance

ρ_{sd} = diffuse reflectance for specular incidence

ρ_{do} = directional reflectance for diffuse incidence

ρ_{dd} = diffuse reflectance for diffuse incidence

The quantities τ_{dd} , ρ_{dd} , τ_{do} , ρ_{do} and τ_{oo} appear twice in Eqs. (2.24). The reason is that no distinction has been made between upward and downward transmittances and between reflectances at the top and at the bottom of the layer for these quantities. It should be noted, however, that the respective quantities are not necessary equal to one another; they will only be so if the layer is homogeneous and if scattering inside the layer is similar for upward and downward radiation.

The interaction with a surface has been described by Eqs. (2.17) and (2.18). Combining these with Eqs. (2.24) enables one to derive the reflectance of a layer that is placed on top of a surface. For this, Eqs. (2.17) and (2.18) are rewritten as

$$E^+(b) = r_{sd} E_s(b) + r_{dd} E^-(b) \quad , \text{ and} \quad (2.25.a)$$

$$\pi L_o^+(b) = r_{so} E_s(b) + r_{do} E^-(b) \quad , \quad (2.25.b)$$

since this interaction takes place at the bottom of the layer. This coupling of a surface and a layer superimposed on it is called the adding method for surface reflectance in four-stream representation. The adding method has been described by various authors, e.g. Van de Hulst (1980) , Cooper *et al.* (1982) and De Haan (1987), and can be applied to problems such as atmospheric scattering in combination with earth surface reflectance, vegetation on a soil, water and sea bottom, etcetera.

The combination of Eqs. (2.24a-d) and (2.25) can be visualized by means of a so-called flux interaction diagram. Fig. 2.2 shows this diagram in order to demonstrate how a scattering layer interacts with the reflectances of the surface and thus produces new resultant reflectances for the top of the layer. In a flux interaction diagram incident fluxes are placed in square boxes and exitent fluxes in circles. Arrows indicate the directions of radiation flow and the reflectances or transmittances beside the arrows are the multiplication factors to be applied in order to calculate the resultant reflectances.

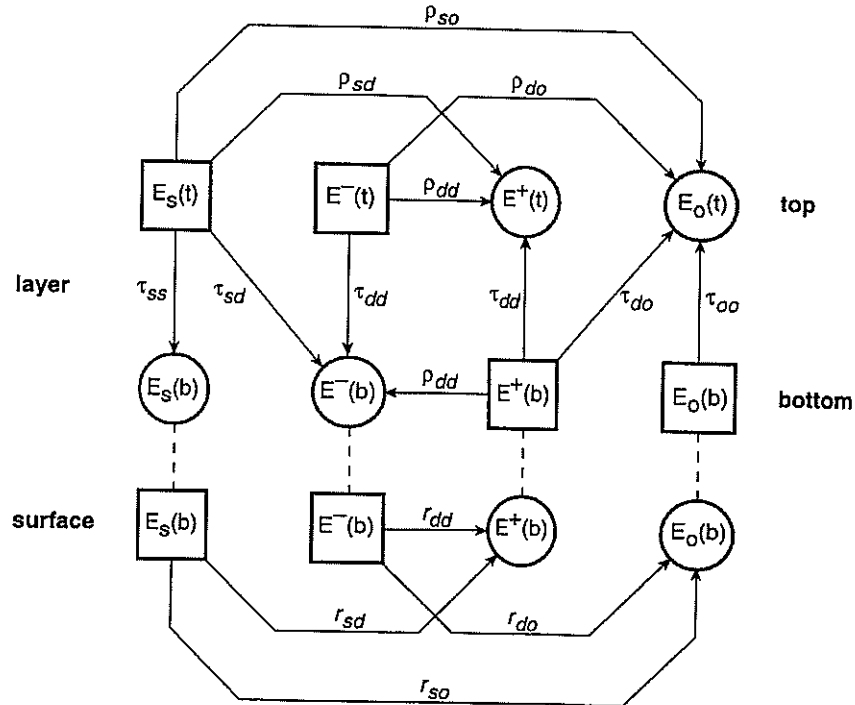


Fig. 2.2 Flux interaction diagram for the combination of a scattering layer and a reflecting surface at the bottom



The radiance in the direction of observation, multiplied by π , πL_o , has been symbolized here by a "flux" E_o . More examples of flux interaction diagrams are shown in chapters 9 and 10 to illustrate the interaction of the atmosphere with the earth's surface.

The resulting reflectances describe the interaction of radiation incident at the top of the layer, with the layer-surface ensemble, as follows:

$$E^+(t) = r_{sd}^* E_s(t) + r_{dd}^* E^-(t) \quad , \text{ and} \quad (2.26.a)$$

$$\pi L_o^+(t) = r_{so}^* E_s(t) + r_{do}^* E^-(t) \quad . \quad (2.26.b)$$

Expressions for r_{so}^* , r_{sd}^* , r_{do}^* , and r_{dd}^* are given in Verhoef (1985a).

2.6 Extinction, absorption and scattering

When a beam of specular radiant flux is incident on a plane-parallel homogeneous layer, the medium inside the layer will partly absorb and partly scatter the intercepted radiation, causing the specular flux density to be attenuated. The attenuation due to this interception is illustrated in Fig. 2.3.

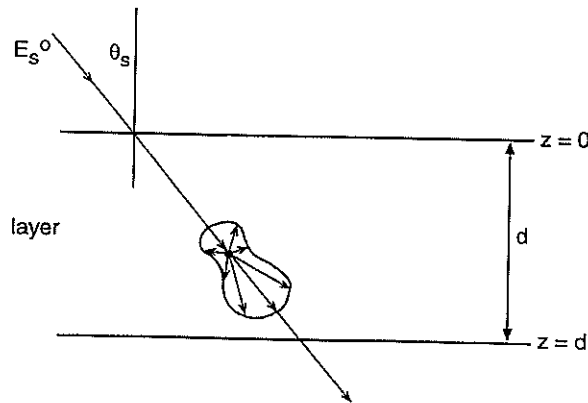


Fig. 2.3 Attenuation of specular flux by scattering

Let E_s^0 be the flux density on a plane perpendicular to the rays; then the flux on a horizontal plane above or at the top of the layer equals



$$E_s = E_s^o \cos \theta_s = E_s^o \mu_s .$$

Inside the layer, E_s is attenuated, and the relative attenuation of E_s per unit metrical depth is called the extinction coefficient, k , so one can write (Beer's Law)

$$dE_s/dz = -kE_s , \quad (2.28)$$

where z is measured in vertically downward direction. The extinction coefficient k depends on the angle of incidence θ_s by

$$k = \beta / \cos \theta_s = \beta / \mu_s , \quad (2.29)$$

because of the increasing path length with increasing θ_s . The quantity β is called the interception coefficient. For so-called isotropic media β is independent of θ_s . This is the case for media consisting of randomly oriented particles, such as most atmospheres.

As β has a dimension of m^{-1} , the product βd is dimensionless and is called the optical thickness, b , of the layer. For layers with vertically varying density one defines the optical thickness as the integral of $\beta(z)$, or

$$b = \int_0^d \beta(z) dz . \quad (2.30)$$

Instead of the metrical depth z one can introduce the optical depth τ by

$$\tau = \int_0^z \beta(z) dz , \text{ and } d\tau/dz = \beta(z) . \quad (2.31)$$

In this way Eq. (2.28) is rewritten as

$$dE_s/d\tau = -(1/\mu_s)E_s . \quad (2.32)$$

The attenuation of diffuse incident flux is derived by replacing E_s with $dE_i = L_i \mu_i d\Omega_i$ in Eq. (2.28):

$$dL_i \mu_i d\Omega_i / dz = -(\beta / \mu_i) L_i \mu_i d\Omega_i .$$



When L_i is constant (isotropic diffuse flux), and both sides are integrated over the upper hemisphere, one obtains $d\pi L_i/dz = -2\pi\beta L_i$, and with $\pi L_i = E^-$, the diffuse downward irradiance, this is written as

$$dE^-/dz = -2\beta E^- \quad (2.33)$$

This means that for diffuse (isotropic) flux the extinction coefficient is twice as great as for specular flux at vertical incidence, or that it is equal to the one found for specular flux at an incidence angle of 60° ($\mu_s = 0.5$). However, these relations are only valid for isotropic media, since β was assumed constant. For non-isotropic media Eq. (2.33) should be replaced by

$$dE^-/dz = -E^- \pi^{-1} \int_0^{2\pi} \int_0^1 \beta(\mu_i, \varphi_i) d\mu_i d\varphi_i \quad (2.34)$$

The scattered radiation can be described by means of the so-called volume scattering function (Kasten & Raschke, 1974) γ' , which is defined by

$$dI_o(\mu_o, \varphi_o) = \gamma'(\mu_s, \varphi_s, \mu_o, \varphi_o) E_s^o(\mu_s, \varphi_s) dV \quad (2.35)$$

where $dI_o(\mu_o, \varphi_o)$ is the intensity in direction (μ_o, φ_o) of the volume dV irradiated by E_s^o from direction (μ_s, φ_s) . Note that E_s^o is the specular flux density on a plane perpendicular to its direction. For a horizontal plane-parallel layer Eq. (2.35) is rewritten as follows by means of the substitutions

$$dI_o = d^2\Phi_o/d\Omega_o \quad ; \quad E_s^o = E_s/\mu_s \quad ; \quad dV = dA dz \quad ,$$

where $d^2\Phi_o$ is the scattered flux in direction (μ_o, φ_o) within a solid angle $d\Omega_o$ and dA is an area element of the layer. Omitting the direction arguments, this gives

$$d^2\Phi_o/d\Omega_o = \gamma' E_s/\mu_s dA dz \quad , \quad \text{or}$$

$$d^2\Phi_o = \gamma' E_s/\mu_s dA dz d\Omega_o \quad .$$

Integration over all scattering directions gives the total scattered flux as



$$d\Phi_o = \int_0^{2\pi} \int_{-1}^1 \gamma' d\mu_o d\varphi_o E_s / \mu_s dA dz \quad .$$

The flux intercepted by volume dV is given by $d\Phi_i = dE_s dA$.
Since $dE_s = (\beta / \mu_s) E_s dz$, the intercepted flux is given by

$$d\Phi_i = (\beta / \mu_s) E_s dA dz \quad .$$

Now $d\Phi_o$ is expressed in $d\Phi_i$ by

$$d\Phi_o = \int_0^{2\pi} \int_{-1}^1 \gamma' d\mu_o d\varphi_o d\Phi_i / \beta \quad .$$

The ratio of total scattered flux to intercepted flux is called the single scattering albedo, ω , and hence is given by

$$\omega = d\Phi_o / d\Phi_i = \frac{1}{\beta} \int_0^{2\pi} \int_{-1}^1 \gamma' d\mu_o d\varphi_o \quad . \quad (2.36)$$

For isotropic scattering (γ' constant) this yields $\omega = 4\pi\gamma'_{is} / \beta$, or $\gamma'_{is} = \omega\beta / (4\pi)$.

The ratio of the actual volume scattering function γ' to the one for isotropic scattering γ'_{is} is called the scattering phase function, $p(\mu_s, \varphi_s, \mu_o, \varphi_o)$ and is given by

$$\begin{aligned} p(\mu_s, \varphi_s, \mu_o, \varphi_o) &= \gamma'(\mu_s, \varphi_s, \mu_o, \varphi_o) 4\pi / (\omega\beta) \\ &= \frac{4\pi\gamma'(\mu_s, \varphi_s, \mu_o, \varphi_o)}{\int_0^{2\pi} \int_{-1}^1 \gamma'(\mu_s, \varphi_s, \mu_o, \varphi_o) d\mu_o d\varphi_o} \quad . \end{aligned} \quad (2.37)$$

From this it also follows that

$$\int_0^{2\pi} \int_{-1}^1 p(\mu_s, \varphi_s, \mu_o, \varphi_o) d\mu_o d\varphi_o = 4\pi \quad .$$

The volume scattering function γ' can be expressed in ω , β and p by

$$\gamma'(\mu_s, \varphi_s, \mu_o, \varphi_o) = \frac{\omega\beta}{4\pi} p(\mu_s, \varphi_s, \mu_o, \varphi_o) \quad . \quad (2.38)$$

An alternative for the volume scattering function γ' is the bidirectional scattering coefficient or area scattering function γ , defined by

$$d\pi L_o = \gamma E_s dz \quad , \quad (2.39)$$

and is closely analogous to the term bidirectional reflectance, defined in section 2.3. The relation between γ and γ' can be derived from Eq. (2.35) by writing the intensity dI_o in terms of the radiance dL_o :

$$dI_o = d^2\Phi_o / d\Omega_o = \frac{d^2\Phi_o}{dA \mu_o d\Omega_o} \mu_o dA = dL_o \mu_o dA \quad .$$

This gives $dL_o \mu_o dA = \gamma' E_s / \mu_s dA dz$, or

$$d\pi L_o = \frac{\pi \gamma'}{\mu_s \mu_o} E_s dz \quad , \text{ so that } \gamma = \frac{\pi \gamma'}{\mu_s \mu_o} \quad . \quad (2.40)$$

Combining this result with Eq. (2.38) gives for the bidirectional scattering coefficient

$$\gamma = \frac{\omega\beta}{4\mu_s \mu_o} p \quad . \quad (2.41)$$

For diffuse incident radiance Eq. (2.35) is modified by replacing the specular incident flux E_s^o with the irradiance on a plane perpendicular to the rays from a solid angle $d\Omega_i$, which is equal to $dE_i^o = L_i d\Omega_i$. This gives

$$d^2I_o = \gamma' L_i d\Omega_i dV = \gamma' L_i d\Omega_i dA dz \quad .$$



Putting $d^2 I_o = \mu_o d^2 L_o dA$ gives

$$\mu_o d^2 L_o dA = \gamma' L_i d\Omega_i dA dz, \text{ or}$$

$$\mu_o d^2 L_o = \gamma' L_i d\Omega_i dz.$$

Integration over all incidence angles, and substitution of γ' from Eq. (2.38) leads to

$$\mu_o dL_o = \frac{\omega}{4\pi} \beta dz \int_{4\pi} p L_i d\Omega_i, \text{ or, with } \beta dz = d\tau,$$

$$\mu_o \frac{dL_o}{d\tau} = \frac{\omega}{4\pi} \int_{4\pi} p L_i d\Omega_i.$$

This expression describes the transfer of radiance from all incidence directions into radiance in a direction (μ_o, φ_o) due to scattering. Incorporation of the extinction of L_o itself gives

$$\mu_o \frac{dL_o}{d\tau} = -L_o + \frac{\omega}{4\pi} \int_{4\pi} p L_i d\Omega_i. \quad (2.42)$$

This is widely known as the radiative transfer equation (RTE) for plane-parallel layers, including absorption and scattering processes.

Eq. (2.42) is restricted to isotropic media since β was treated as a constant in its derivation.

For non-isotropic media the concept of optical depth τ loses somewhat of its meaning and in this case the RTE is preferably written in terms of the metrical depth z as

$$\frac{dL_o}{dz} = -k L_o + \pi^{-1} \int_{4\pi} \gamma L_i \mu_i d\Omega_i. \quad (2.43)$$

where k is the extinction coefficient and γ the bidirectional scattering coefficient.

In the literature on radiative transfer in vegetation canopies the function $\beta(\mu)$, or even $\beta(\mu, \varphi)$, barring a proportionality constant, is called the G-function (e.g. Ross & Nilson, 1975). Basically it describes how the inter-



cepted flux depends on the direction of incidence. In the Suits model (Suits, 1972) and the SAIL model (Verhoef, 1984) this function is not identified as such, but use is made of the extinction coefficients $k = k(\mu_s)$ and $K = k(\mu_o)$. In essence, these models are based on the equations

$$dE_s/dz = -kE_s \quad ,$$

$$d\pi L_o/dz = w(\mu_s, \varphi_s, \mu_o, \varphi_o)E_s + \int_{4\pi} w(\mu_i, \varphi_i, \mu_o, \varphi_o)L_i \mu_i d\Omega_i - K\pi L_o \quad ,$$

where the first describes the extinction of specular (solar) flux, and the second is the actual RTE, with w as the bidirectional scattering coefficient and K as the extinction coefficient in the scattering direction.

So far this introduction of fundamental radiometric quantities. In the following chapters these quantities and the relations between them will frequently be applied for the physical description of radiative transfer in different media such as the terrestrial atmosphere and vegetation canopies.



3 CLASSIFICATION OF RADIATIVE TRANSFER MODELS

In this chapter radiative transfer models are discussed on the basis of a number of features, which makes it possible to distinguish existing models from one another and to place the models discussed in this thesis in a proper context. The features considered here are :

- type of medium
- scattering behaviour
- representation of radiative transfer equation
- method of solution

After a discussion of these features, a section is devoted to the subject of coupling methods, which is important for models consisting of a combination of different types of medium in different layers.

3.1 Type of medium

For the application in remote sensing problems, radiative transfer models have been developed for a wide variety of media, a list of which follows below :

- air
- aerosols, clouds
- water bodies
- snow
- plant leaves
- vegetation canopies
- forests

From a physical point of view, scattering media can also be classified according to their properties with respect to the interception of incident radiation. This leads to the following categories :

- isotropic media
- azimuthally isotropic media
- anisotropic media



The distinction between these types of interception is very important as it largely determines the methods of solution that can be chosen. For isotropic media the interception of radiation is independent of the direction of incidence. This is the case for collections of spherical particles and for collections of non-spherical particles which have a random orientation. Examples are the atmosphere, water, and homogeneous vegetation canopies with a spherical leaf angle distribution.

For azimuthally isotropic media the interception of incident radiant flux is independent of the azimuth angle, but it does depend on the zenith angle. In general this is the case for media in which e.g. the force of gravity imposes an orientation of the elements which is random in azimuth but non-spherical in the zenith direction. Examples are homogeneous vegetation canopies and snowflakes in the atmosphere. Anisotropic media intercept radiation depending on both the zenith and the azimuth angle of incidence. Vegetation canopies with a heliotropic leaf orientation distribution and those exhibiting a clear row structure are examples of this category.

3.2 Scattering behaviour

For the scattering of incident radiation by collections of elementary particles physical descriptions have been derived by various authors. In atmospheric scattering those for air and aerosol particles are known as Rayleigh- and Mie-scattering, respectively. Polarization effects are also included in these theories. For the explanation of rainbows and similar phenomena use is made of theories based on geometric optics.

Also the interception and scattering of light by leaf layers has been described in physical terms. Here, usually one assumes that the light intercepted by a leaf is scattered in a bi-Lambertian manner, meaning that the reflected and transmitted light are both isotropic, but not necessarily equal. In some more sophisticated theories, also a specular component of the leaf reflectance is included, as well as polarization effects.

3.3 Representation of the radiative transfer equation

The radiative transfer equation (RTE) forms the basis of each optical model that allows to predict the reflected and transmitted radiation from the object's



properties and the incident radiation. The RTE is a so-called integro-differential equation and describes the differential change of the radiance in a certain direction due to the combined effect of interception in that direction and the effect of scattered incident flux integrated over all angles of incidence. There exists no analytical solution to this equation, so that one always has to accept some form of approximation.

By representation of the RTE is meant a modified RTE which is numerically solvable, and which is somehow adapted to the type of medium and scattering at hand. In this respect mostly one-dimensional plane-parallel media are considered here, as these form the majority of existing models. In such media the most important metric quantities are the metrical depth z , and the zenith and azimuth angles. As to the representation of the RTE one can now distinguish between some types and levels of discretization that are usually applied :

- angular segments
- discrete angles
- discrete layers vs. continuous z

The latter discrimination results into a set of difference equations versus a set of differential equations representing the RTE, thus leading also to a purely numerical, or to an analytical solution. As to the angular discretizations, for angular segments always a fixed tessellation of the sphere of directions is used, whereas in the case of discrete angles the choice of directions can be either fixed or adaptive.

In many existing models the above mentioned discretizations apply only to the description of the diffuse flux field. For example, the so-called single scattering contribution can be calculated analytically when two discrete directions, namely the direction of the sun and the direction of viewing, are given.

Finally it should be mentioned that for media in which the size of the scattering elements becomes significant, a description of radiation transfer by means of just an RTE is incomplete, since in that case shadowing effects must also be accounted for. In particular, this situation is found in models of forests and of heterogeneous vegetation in general. In this case one usually combines a description in terms of an RTE with one based on geometric optics.



3.4 Methods of solution

The methods of solution to be discussed in this section are the ones most frequently found in the literature on remote sensing and optics. These are

- Monte Carlo method or ray-tracing
- adding method
- doubling method
- successive orders of scattering approximation (SOSA)
- analytical solution by means of eigenvectors
- discrete ordinates method (DOM)

Ray-tracing

In this method photons are numerically 'traced' through the medium until they are either absorbed or escape from the medium. By counting the photons that manage to escape one can determine the reflectance and the transmittance. Each change of direction inside the medium is governed by the probability of scattering in the exitence direction given the incidence direction and the scattering properties of the medium. What actually happens to a photon is determined by drawing random numbers.

A great advantage of this method is that it can be applied to almost any scattering medium, so that the level of realism can be very high. For instance, heterogeneous media such as forest stands are difficult to model realistically without incorporating their 3-D structure and a ray-tracing approach can be a good solution in this case. However, a disadvantage is the fact that this method is computationally very demanding, since for the determination of e.g. the bidirectional reflectance thousands of photons escaping in the exitence direction must be counted. Also, the relation between the reflected and transmitted radiation and the properties of the medium becomes almost as complex as in reality. A useful application of this method is the testing and validation of simpler models.

Adding method

The adding method is applied to discrete layers of finite optical thickness, whereas the angular discretization can be arbitrary. This method is based on the finding that the optical properties of a stack of two layers can be com-



puted from the optical properties of the two separate layers. These separate layers do not have to be optically thin, and therefore the adding algorithm can also be applied successfully in the computation of the optical properties of vertically heterogeneous media. Normally however, this method is applied in such a way that, starting with one optically thin layer, identical layers are added successively until the required optical thickness has been reached.

As a by-product, the adding method gives also the vertical flux profile. In practice, this is the only reason to apply the method to homogeneous media, as when the vertical flux profile is not required, the much more efficient doubling method, discussed hereafter, can be used.

Doubling method

The doubling method starts with an optically thin layer, and uses the adding equations in order to find the optical properties of a stack of two of these layers. This doubling step is repeated until the required optical thickness has been reached. For vertically homogeneous media this method is much more efficient than the adding method, since for instance a layer consisting of 1024 sublayers would require 1023 adding steps and only 10 doubling steps.

Successive orders of scattering approximation (SOSA)

In this method the medium is divided into a finite number of sublayers and the diffuse flux field in all sublayers is computed iteratively. This is done by starting with the so-called zero-order flux, i.e. the flux field resulting from interception alone (no scattering). Next this flux field interacts with the medium to produce the first-order (singly scattered) flux field, which interacts with the medium to produce the second-order field, etcetera. This continues until successive approximations no longer influence the final result, i.e. the total flux field at all levels. The reflectance and the transmittance of the medium can finally be found from the upward flux at the top layer and the downward flux at the bottom layer, respectively. This method is conceptually elegant and particularly useful if the internal flux field has to be determined. A disadvantage is that in case of low absorption the iteration can be very slow, for instance in the atmosphere and in vegetation canopies for near infrared radiation.



Analytical solution by means of eigenvectors

In the context of radiative transfer, an eigenvector is a set of weights applied in a linear combination of the fluxes in all discrete directions considered in the representation of the RTE. For each linear combination the radiative transfer equation acts in a special way, namely that only extinction in the z -direction has to be considered, which means that each flux pattern associated with an eigenvector persists through the entire layer and that its strength decreases or increases exponentially as a function of z . Each eigenvector has its own extinction coefficient, called eigenvalue. For the solution, first the incident flux field is decomposed into its eigenvectors, after which the exponential extinction is applied in order to find the intensities of the eigenvectors at the top and the bottom of the layer. Finally, these are transformed back to original fluxes, from which reflectance and transmittance can be determined.

This method is particularly useful if calculations have to be carried out for many values of the optical thickness if everything else is constant, since the eigenvectors and eigenvalues in that case are also constants.

A disadvantage of the method is that the mathematical formulation cannot be translated directly into safe computer code, as mathematical singularities can lead to serious numerical problems. However, if these are properly taken care of, a computationally very efficient method is the result.

Discrete ordinates method (DOM)

The discrete ordinates method, as the name suggests, employs a set of discrete directions in order to represent the RTE. However, in a more strict sense, this method refers to the situation of an isotropic medium, in which case the choice of directions is optimized on the basis of Gaussian integration. Here the scattering phase function is expanded in a set of Legendre polynomials and the azimuthal flux field in a Fourier series. If no polarization is included, then the latter reduces to a cosine-series. As to the Gaussian quadrature scheme, both full-range and double half-range (in zenith direction) variations have been reported. In the latter case each hemisphere is treated separately, which gives the advantage that the discontinuous behaviour of the flux at the horizon is more easily accommodated.

The final result is a linear system of differential equations, which can be solved analytically by means of the method above, or they are turned into a



set of difference equations after which adding, doubling or SOSA can be applied for a solution.

3.5 Coupling methods

For the calculation of the optical properties of an ensemble of different layers, or of the reflectance of a surface covered by one or more scattering layers, it is necessary to use a so-called coupling algorithm.

Besides the coupling algorithm itself, it is also important to consider what restrictions are imposed on the boundary equations when applying a particular method of solution.

As to the coupling methods, the most frequently used are the adding method and the boundary equation method. The adding method has been explained already above and can be applied to turbid media, i.e. media in which the size of the scattering particles is negligible. In this respect, the boundary equations method is more general, since in that case this restriction is not required. A disadvantage of the boundary equations method is that a system of linear equations has to be solved. This system of equations is the result of the boundary conditions at the top and the bottom of the ensemble of layers and of the required continuity of the fluxes at the interfaces between layers.

Restrictions on the boundary equations are particularly important when the types of media in the different layers are different or when the reflectance of a surface at the bottom of an ensemble of layers has properties which cannot easily be accommodated by the representation of fluxes in the layers above. For instance, in the combination water surface - atmosphere the surface generates upward specular flux and if this kind of flux is not included in the atmospheric layers above, then the boundary equations at the bottom are restricted, so that less realistic results would be obtained. Also, a DOM-based atmospheric model cannot easily be combined with a surface vegetation reflection model as the latter must be based on a different method, because the vegetation medium is in general non-isotropic.



3.6 Conclusion

Radiative transfer models vary widely according to the type of medium modelled, the physical laws of scattering behaviour, the representation of radiative transfer by means of mathematical expressions, and the various methods of solution that can be applied. Except for the former two characteristics, these features are relatively independent of one another, so that in practice there is a great variety of possible model implementations one can choose from. In the next chapter a number of radiative transfer models for vegetation and the atmosphere are discussed on the basis of these characteristics, thus illustrating combinations of features included in models that actually have been implemented by various investigators.



4 MODELS FOR VEGETATION AND THE ATMOSPHERE

This chapter gives a brief review of existing radiative transfer models for vegetation canopies and atmospheric layers. The features discussed in chapter 3 will serve to some extent in order to classify the models into different categories.

4.1 Vegetation models

Radiative transfer models for vegetation canopies have initially been developed for the study of photosynthesis in plant stands. Stimulated by optical remote sensing techniques, later also models were developed in order to obtain a better insight in the interaction between incident radiation and vegetation canopies, giving support to the interpretation of remotely sensed images of crops. An excellent and reasonably up-to-date overview of vegetation canopy reflectance models is given by N.S. Goel (1988). This 212 pages thick review article contains about 200 references and discusses about 40 different existing models. The models discussed in this chapter form only a small subset of these, and were selected on the basis of lines of development and representativeness of certain types of model in terms of the features discussed in the previous chapter.

4.1.1 Discrete models

This class of models is based on the division of the canopy into a finite number of layers, for which radiative transfer is described by means of a set of difference equations. One of the first of these is the model of De Wit (1965). Goudriaan (1977) has extended and improved this model (mainly with respect to the method of solution). Since the calculation of photosynthesis was the primary interest, the model employs a rather coarse angular division, namely 9 zenith zones of 10 degrees each, and no azimuthal division. Also, leaves are assumed to have equal reflectance and transmittance, as expressed by the use of only one constant scattering coefficient. For photosynthesis calculations this may be no problem, it is for reflectance calculations, especially in the visible part of the spectrum, where the relative difference between leaf reflectance and transmittance can be substantial. Goudriaan uses a method of solution which is iterative in nature and which



in essence does not differ from the SOSA method. In Bunnik (1978) results of Goudriaan's model are compared to those of the Suits model, discussed later.

The model of Cooper, Smith & Pitts (1982) is very similar to Goudriaan's model, except that the adding method of Van de Hulst (1980) is used as the method of solution, and that single leaf reflectance and transmittance can be unequal.

Den Dulk (1989) in his models applies the original idea of using discrete directions chosen from a polyhedron consisting of hexagons and pentagons ('soccer ball'). This gives 92 directions in total with a mean angular difference of 24 degrees. The diffuse flux field as well as the leaf orientation distribution can be expressed in these 92 directions, so that even azimuthal symmetry is not a requirement. Den Dulk calls his models TURTLE and HARE, the first one of which uses the adding method, so that it can also provide the internal flux profile, and the second one uses the much faster doubling method.

4.1.2 Analytical models

By analytical models is meant the class of models which is based on a continuous vertical dimension, so that radiative transfer is described by means of a set of differential equations. The simplest of these is formed by the Kubelka-Munk (1931) theory, which has been applied to all sorts of media, for instance layers of paint, plant leaves, turbid water, etcetera.

Because this theory is based on two semi-isotropic up- and downward diffuse fluxes it is referred to as a two-stream approximation. In terms of angular discretization this theory uses two segments, namely the upper and lower hemisphere. The original KM-equations are given by

$$\begin{aligned} dI/dz &= -(k+s)I + sJ, \quad \text{and} \\ dJ/dz &= (k+s)J - sI, \end{aligned} \tag{4.1}$$

where I = downward flux

J = upward flux

k = absorption coefficient

s = (back) scattering coefficient

z = metrical depth



Fig. 4.1 illustrates this interaction of fluxes inside a turbid plane-parallel medium.

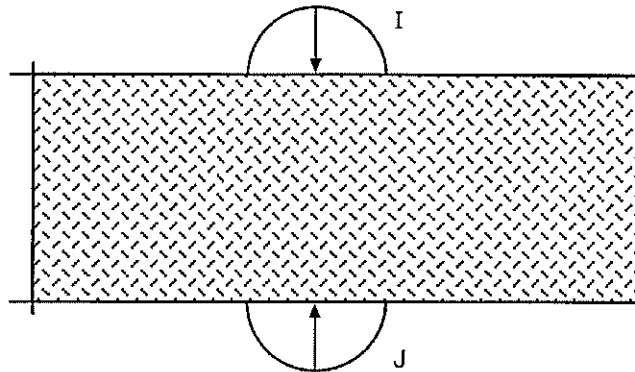


Fig. 4.1 Diffuse flux interactions according to Kubelka-Munk theory

The system of coupled linear differential equations can easily be solved analytically. For solutions, see for instance Bunnik (1978) or Den Dulk (1989).

Eqs. (4.1) describe the increase of flux due to backscattering of flux from the opposite direction and the decrease caused by loss of the flux in the same direction due to absorption and backscatter. The combination $(k + s)$ is the so-called attenuation coefficient.

In an alternative notation, which is more compliant with the one used elsewhere in this manuscript, these equations are written as

$$\begin{aligned} dE^-/dz &= -a E^- + \sigma E^+ \quad , \quad \text{and} \\ dE^+/dz &= -\sigma E^- + a E^+ \quad , \end{aligned} \tag{4.2}$$

where E^- = downward flux

E^+ = upward flux

a = attenuation coefficient

σ = backscattering coefficient

In order to include also the effects of a specular source of flux like the sun, the above system of equations was extended by Duntley (1942). The so-called Duntley equations are given by



$$\begin{aligned}
 dE_s/dz &= -k E_s, \\
 dE^-/dz &= s' E_s - a E^- + \sigma E^+, \\
 dE^+/dz &= -s E_s - \sigma E^- + a E^+,
 \end{aligned} \tag{4.3}$$

where E_s = specular flux

k = extinction coefficient for specular flux

s' = forward scatter coefficient for specular flux

s = backscatter coefficient for specular flux

In the vegetation canopy reflectance model of Allen, Gayle & Richardson (1970) these equations are applied to simulate the hemispherical reflectance as a function of the solar elevation during the day. As the model has five parameters, five independent measurements must be carried out at least for an empirical fit.

The model of Suits (1972) also uses the Duntley equations, but now the coefficients are directly expressed in biophysical parameters of the canopy. In addition to that, Suits also introduces four extra coefficients (an extinction coefficient and three scattering coefficients) in order to estimate the radiance at the top of the canopy for calculation of the bidirectional reflectance. A disadvantage of the Suits model is that the coefficients are only defined for horizontal and vertical leaves, so that for an actual leaf inclination distribution the oblique leaves must be projected on horizontal and vertical planes in order to obtain the total horizontal and vertical leaf area indices, H and V . For the vertical projections a random azimuth distribution is assumed.

The SAIL model (Verhoef, 1984), from Scattering by Arbitrarily Inclined Leaves, is a refinement of the Suits model in that the coefficients can be computed for any leaf inclination, while still a random leaf azimuth distribution is assumed. This model is further discussed in chapter 7.

Incorporation of the so-called hot spot effect in a manner suggested by Kuusk (1985) has resulted in a variant called SAILH, which is discussed in chapter 8 of this thesis. In SAILH the size of the leaves and the associated shadowing effects are taken into account for the calculation of the single scattering contribution to the bidirectional reflectance. All other contributions are computed as before. Except for the hot spot effect in SAILH, all of the above models are solved analytically.

In Rosema *et al.* (1991) a variant called FLSAIL is introduced which is not



solved analytically, but by means of the doubling method. This model was intended for the simulation of canopy fluorescence and since the combined fluorescence and scattering processes are expressed by eight differential equations, a purely analytical solution would become too tedious. By this example it is also demonstrated that analytical and discrete models are sometimes less different from one another than might be expected.

The analytical models discussed so far describe the diffuse flux field by means of the very crude angular segmentation in only two hemispheres. The extensions towards the incorporation of specular solar flux by Duntley and of radiance in the direction of observation by Suits have enormously increased the applicability of the KM-theory, but the effects of deviations of the diffuse flux from semi-isotropic are still relatively unknown. Den Dulk (1989) has compared SAIL with his model and this revealed virtually no difference in output, so for vegetation canopies this approximation seems to be justified. An extension of the basic Kubelka-Munk theory towards a matrix-vector formulation in order to accommodate a finer angular segmentation has been introduced by Chen (1985). He uses nine zenith zones and 36 azimuth sectors in each hemisphere, which leads to matrices of dimension 324×324 to describe diffuse scattering. Although he suggests ways to reduce computation time for the case of azimuthally isotropic canopies, his solution seems more involved than application of an azimuthal cosine-transform, which would result in only nineteen 9×9 matrices (see chapter 5 of this thesis).

This section is concluded with the differential equations representing the so-called four-stream theory which is used in the Suits and SAIL models. In a notation slightly different from the one in Verhoef (1985a) these read

$$dE_s/dz = -k E_s, \quad (4.4.a)$$

$$dE^-/dz = s' E_s - a E^- + \sigma E^+ \quad (4.4.b)$$

$$dE^+/dz = -s E_s - \sigma E^- + a E^+ \quad (4.4.c)$$

$$dE_o^+/dz = -w E_s - v E^- - v' E^+ + K E_o^+ \quad (4.4.d)$$

$$dE_o^-/dz = w' E_s + v' E^- + v E^+ - K E_o^- \quad (4.4.e)$$

where $E_o = \pi$ times the radiance in the direction of view

K = extinction coefficient in the viewing direction

v = directional backscatter coefficient for diffuse incidence



ν' = directional forward scatter coefficient for diffuse incidence
 w, w' = bidirectional scattering coefficients for downward, upward view

The last equation is included here for completeness, since it describes the generation of downward directed radiance. It is not necessary for calculation of the bidirectional reflectance of canopies. Eqs. (4.4.d-e) were introduced in Verhoef (1985a) in order to better comply with formulations of radiative transfer in atmospheres. Suits (1972) uses the concept of 'direct line of sight' for derivation of the radiance at the top of the canopy. In essence, he applies two differential equations, namely

$$\begin{aligned} dF_o/dz &= -KF_o, \\ dE_o(t)/dz &= -(wE_s + \nu E^- + \nu' E^+) F_o, \end{aligned} \quad (4.5)$$

where F_o = the fraction of a layer seen by direct line of sight
 $E_o(t)$ = π times the radiance at the canopy top

However, it can be shown that F_o is a redundant quantity and that Eq. (4.4.d) can be used instead, which is more general since it allows to determine the radiance at any level within the canopy, not just at the top.

4.1.3 Other models

Models not discussed so far are those that apply other representations of radiative transfer than the ones outlined above for homogeneous canopies, or that apply a mixture of solution techniques for heterogeneous vegetation.

Ross & Nilson (1975) describe the interaction of light with a canopy in very general terms using the radiative transfer equation of chapter 2. They employ a scattering phase function which is not limited to isotropic media and they include non-Lambertian reflection and transmission of leaves in their description. Their theory has been the starting point for various refinements, such as the incorporation of the hot spot effect (Nilson & Kuusk, 1989), and for alternative methods of solution like the Monte Carlo technique (Ross & Marshak, 1984).

Myneni, Asrar & Kanemasu (1987) also follow this approach and use the SOSA method for the solution.



A number of models also are based on descriptions in terms of the radiative transfer equation, but assume an isotropic medium in order to enable the use of well-established solution methods like the DOM. In such a medium the scattering phase function depends only on the angle between the incidence direction and the exitence direction. Examples are the models of Weinman & Guetter (1972), Gerstl & Zardecki (1985), and Camillo (1987). However, since only homogeneous canopies with a spherical leaf angle distribution really fulfil this condition, the general applicability of these models seems questionable.

Contrary to the so-called one-dimensional models discussed so far, heterogeneous vegetation showing density variations due to clumpiness, rows, etcetera, should be treated as a three-dimensional medium. One of the first attempts to model a row crop was reported in Verhoef & Bunnik (1976). This model was based on the Suits model, but calculated the contribution from singly scattered sunlight on the basis of an abstraction of the rows shaped as rectangular prisms. Suits (1983) relaxed this rather crude abstraction by allowing more general periodic density variations and introduced numerical calculations for the estimation of the diffuse fluxes. The results of both models showed that in, or close to, the row direction the bidirectional reflectance substantially differs from the one of the equivalent homogeneous canopy, especially in the red part of the spectrum. Extensions of the SAIL model for heterogeneous vegetation have been created by Goel & Grier (1986). Their model called TRIM (Three-dimensional Radiation Interaction Model) uses regularly spaced elliptical subcanopies in order to represent the density modulation. Although this model is internally very complex, the number of input parameters has been kept to a minimum in order to retain good possibilities for parameter estimation by model inversion techniques. The model called BIGAR of Norman & Welles (1983) is even more general than TRIM in that it allows to specify subcanopies shaped as ellipsoids, cylinders, cones, etcetera, in whatever spatial arrangement desired.

A rather flexible concept for modelling of heterogeneous vegetation was introduced in the model of Kimes & Kirchner (1982). Here the 3-D canopy is discretized as a volume consisting of cubic cells, and also the angular space is segmented into zenith zones and azimuthal sectors. The solution method is iterative in nature and is partly inspired by the SOSA method and by ray-tracing techniques.



4.2 Atmospheric models

A detailed review of the massive literature on radiative transfer models for the terrestrial atmosphere falls outside the scope of this thesis. Instead, only a few models are discussed here because of their connection with the models described in chapters 5 and 9.

The model of R.E. Turner (1973) is an example of a relatively simple approach to the problem of atmospheric scattering in relation with the analysis of optical remote sensing data. The model is rather detailed with respect to the height profiles of optical thicknesses of various atmospheric constituents and further is heavily based on the finding that aerosol scattering is highly anisotropic, with a pronounced forward scattering peak. In the model the single scattering contribution is described by means of the actual scattering phase function, so that the calculation of this part is exact, but for the multiple scattering contribution a so-called double delta approximation is applied, which simplifies the radiative transfer equation to a pair of linear differential equations that can be solved analytically. The double delta approximation simplifies the multiple scattering of aerosols by the assumption that one fraction of the incident radiation is scattered exactly forward and another fraction is scattered exactly in backward direction. This means that no radiation escapes out of the beam and for each direction the transfer of radiation can be described by two differential equations of a type similar to the Kubelka-Munk equations, but with the singly scattered fluxes in that beam added as source terms. This approach leads to an unrealistically high peak in the anti-solar direction, which can be explained by a "reflection" of the forward peak of singly scattered direct solar flux. In reality the forward scattering peak of singly scattered solar flux is spread over all backward directions and therefore the upward flux is much more uniform than predicted by this model. Nevertheless, the idea of simplifying the scattering behaviour of aerosols to a forward fraction and a backward fraction has inspired the author in the design of the four-stream atmospheric model described in chapter 9. Here it is assumed that the multiply scattered flux field can be approximated by means of two semi-isotropic hemispherical fluxes, just like is done in the SAIL model. The associated scattering coefficients of this model are based on the assumption of a fixed division of the incident flux in a forward scattered fraction and a backward scattered fraction, as in Turner's model.

An example of a model based on the discrete ordinates method (DOM) is the



one of K. Stamnes *et al.* (1988). This model is quite general in scope as it incorporates also the thermal infrared emission of the atmosphere and the earth's surface. Polarization is not included. The angular integration of the radiative transfer equation is carried out by means of a double half-range Gaussian quadrature in the zenith angle domain, i.e. both hemispheres are integrated separately. This complies better with the discontinuity of the radiance at the transition from upper hemisphere to lower hemisphere than does a full-range Gaussian integration. By expanding the scattering phase function into Legendre polynomials and the azimuthal radiance in a Fourier cosine series, the zenith angle dependence and the azimuthal dependence are uncoupled, and the result is expressed by M systems of N linear differential equations, where M is the number of azimuthal cosine terms and N the number of Legendre polynomials. Each system is solved analytically by means of eigenvector decomposition and in this respect there is much similarity with the $(N+2)$ -stream model described in the next chapter. Vertical heterogeneity of the atmosphere is accommodated for by a stratification into layers that can be considered homogeneous. Coupling of the different layers and the reflectance of the lower boundary is carried out by means of the boundary equation method, not the adding method. The bidirectional reflectance of the earth's surface is incorporated by assuming that it can be expanded in Legendre polynomials. However, this implies a restriction to media for which the bidirectional reflectance depends only on the angular difference between the directions of the incident and the reflected radiation, which for vegetation canopies is not very realistic.

4.3 Discussion

The radiative transfer models for vegetation canopies and atmospheres discussed in the previous sections illustrate that various approaches can be followed in order to obtain a model that forms a reasonable compromise between computational complexity and realism.

Computational complexity can either be the result of incorporating a very large number of effects, like heterogeneity, complicated scattering behaviour, azimuthal anisotropy, etcetera, all with their associated parameters, or it may be caused by using a more accurate calculation method without introducing any extra parameters or effects. The first category of models leads to greater realism, but also to the necessity of specifying all the additional parameters,



which sometimes may be hard to obtain. The second category gives greater realism just because the calculations are more accurate, probably at the cost of longer computation times. The input parameters and the effects incorporated remain the same.

A vegetation model like SAIL could be improved in both directions, since many effects like leaf gloss, crop row effects and the finite leaf size are not incorporated, and the method of solution is only approximative in nature. However, the most important effects are included in SAIL and a greater accuracy would be of doubtful value if the uncertainty about some of the input parameters is much greater, which is usually the case. More details on the SAIL model are given in chapter 7. One significant extension of SAIL that has been effectuated is the incorporation of the hot spot effect, which is caused by the finite leaf size. This extension, leading to the so-called SAILH model, is discussed in chapter 8.

An example of extension of a model in the direction of higher accuracy without introducing additional input parameters is formed by the so-called $(N+2)$ -stream model, discussed in the next chapter. This model is a refinement of a similar four-stream theory, which is applicable to vegetation canopies (SAIL) as well as to the atmosphere.



5 (N+2)-STREAM THEORY

5.1 Introduction

Bidirectional reflectance modelling of vegetation canopies provides a lot of insight in the relations between biophysical crop parameters and the signals detected by remote sensing equipment on board aircraft or satellites. However, these sensors do not detect the bidirectional reflectance as such, but rather the radiance at some level above the ground, and although this radiance is proportional to the directional reflectance of the target, it is also influenced by the atmosphere between target and sensor.

In principle the atmospheric influence could be simulated by means of a separate radiative transfer model for the atmosphere, but this would be complicated by the fact that the atmosphere and the earth's surface interact, so that several iterations would be necessary. This can be illustrated by means of the ratio of diffuse/direct incident flux, which is a required input in most canopy reflectance models. This ratio strongly depends on the atmospheric turbidity, but because of the repeated reflections of radiation between the bottom of the atmosphere and the surface, it depends also on the reflectivity of this surface, which can be estimated using the canopy reflectance model. In other words, the diffuse downward flux required as input by the canopy reflectance model, depends partly on the output of this model.

In chapter 3 the adding method was mentioned as a tool to compute the optical properties of a pile of different layers from those of the individual layers, and this technique can also be used to couple a canopy reflectance model with an atmospheric model. This would eliminate the need to specify the diffuse downward flux as input to the canopy reflectance model, and the surface reflectance as input to the atmospheric model, as both would be computed internally in the integrated model. However, a requirement for application of the adding method is that at the interface between two layers the radiant fluxes must be compatible, or should be made compatible. This means that, for example, if the SAIL model is used as the canopy reflectance model, which is based on semi-isotropic, hemispherical diffuse fluxes, then the atmospheric model should also produce and accept similar fluxes at the interface. If this is not the case, then conversions are inevitable, and these should be integrated in the computer program.

A first attempt to couple the SAIL model with an atmospheric model has been presented in Verhoef (1985b). Here the atmospheric model was based



on the same four-stream differential equations as where the SAIL model is based on, with the extinction and scattering coefficients expressed in atmospheric quantities like scattering phase function, optical thickness and single scattering albedo. As in both models the same fluxes are used, there are no compatibility problems in applying the adding method in this case. However, although the four-stream approach has been reasonably successful for vegetation canopies, it may not be sufficiently accurate for atmospheres, especially when considering the highly anisotropic scattering behaviour of aerosols.

Most numerically accurate atmospheric radiative transfer models make use of the discrete ordinates method (DOM), which is computationally efficient, but assumes that the medium is isotropic (constant interception coefficient), which implies that the phase function can be expanded in Legendre-polynomials, and the dependencies in zenith and azimuth can be decoupled by means of an azimuthal Fourier-transform. Except for homogeneous leaf canopies having a spherical leaf angle distribution, a vegetation layer is not isotropic, at least not in zenith direction. Therefore, the discrete ordinates method cannot be applied in the same way to leaf canopies as it is applied to atmospheres.

Another possibility is to take as a starting point the four-stream vegetation model and to refine this with respect to the angular resolution, which with only two hemispherical fluxes is rather coarse. In that case the refined model can also be applied to the atmosphere, and the adding method can be applied for the coupling of both models without any compatibility problems. In order to improve the numerical accuracy of the four-stream model for atmospheres, the four-stream concept has been generalized by the author to a so-called $(N+2)$ -stream concept. In the $(N+2)$ -stream model the two semi-isotropic diffuse fluxes are replaced by N locally isotropic fluxes originating from N segments of the sphere of directions. The first results of this model have been obtained for $N = 72$, and were based on a direct generalization of the corresponding four-stream equations (Verhoef, 1985a). Although this was notationally attractive, as in most formulas of four-stream theory scalar quantities only had to be replaced by vectors and matrices as their $(N+2)$ -stream counterparts (Verhoef, 1988), numerical problems were encountered, and careful analysis revealed that these were all related to conditions under which eigenvalues were almost or exactly equal. In $(N+2)$ -stream theory radiative transfer is described by a set of $N + 2$ linear differential equations, of which the general solution can be expressed in



linear combinations of exponential functions. The coefficients of these exponential functions are the eigenvalues of the matrix containing all extinction and scattering coefficients. Two eigenvalues of this matrix are equal to the extinction coefficients associated with direct solar flux and with radiance in the direction of view, respectively. These are called k and K . The other eigenvalues are obtained from the diffuse scattering matrix, and because of mirror symmetry between the directions upward and downward, these occur in $N/2$ positive-negative pairs. Now it is a mathematical law that a system of $N + 2$ linear differential equations must have $N + 2$ linearly independent solutions. However, this law is violated if two eigenvalues are equal, and in that case a particular solution, which is not purely exponential, must be added. This problem can occur if k or K becomes equal to one of the positive eigenvalues associated with the diffuse scattering matrix, or if the smallest eigenvalue equals zero, which occurs in the case of conservative scattering (no absorption). Of course the problem of equal eigenvalues can also be encountered when $N = 2$, i.e. in the four-stream atmospheric model or in the SAIL model. However, in practice this situation did not manifest in the results, probably because exact equality of eigenvalues is rarely reached. Nevertheless, as the results showed, for greater values of N , e.g. $N = 72$, the numerical results appear to become more sensitive to this.

In a computer program for calculation of radiative transfer in a layer the case of precisely equal eigenvalues can easily be trapped, and a suitable subroutine for handling this case can be called. But it would not be clear what to do in case the difference between two eigenvalues is very small. Strictly, this situation should be handled as the 'normal' case, as the eigenvalues are not equal, but a detailed analysis has shown that in this case numerical problems can arise because some terms contain the difference of the eigenvalues in the denominator, so that these terms become very large. Although this may still be mathematically correct, differences between very large terms can become very inaccurate in a computer.

Because of the above mentioned numerical problems, the solution of the $(N+2)$ -stream model has been completely reformulated recently (1994). In this revised version all potential numerical problems are avoided, not only the case of equal eigenvalues, but also those associated with extremely high optical thicknesses. Furthermore, much effort has been put in checking the physical validity of the formulas, for instance by deriving approximations for small optical thickness, and by verifying the so-called reciprocity relations. For the bidirectional reflectance and transmittance of a layer this has resulted



in new expressions which are completely symmetric with respect to the directions of specular incidence and of observation, so that the reciprocity relations are automatically fulfilled.

This chapter first discusses the subject of tessellation, i.e. the division of the sphere of directions into angular segments for the description of the diffuse flux field (section 5.2). In section 5.3 it is shown how the $(N+2)$ -stream differential equations are derived from the radiative transfer equation, and some of the properties of the elements of the coefficient-matrix will be discussed. The analytical solution of the system of differential equations by means of an azimuthal discrete cosine transform (DCT) and further decomposition into eigenvectors is discussed in section 5.4 and Appendix A. Special attention is paid to the measures taken in order to avoid numerical problems associated with (nearly) equal eigenvalues. Section 5.6 and Appendix B deal with the mathematical validation of the model by analyzing its behaviour for small and infinite optical thickness and by verifying the so-called reciprocity relations.

5.2 Tessellation

As mentioned in section 2.3, the radiative transfer equation (RTE), i.e. Eq. (2.42) of section 2.6, cannot be solved analytically because it involves an integration over all incidence directions, and for arbitrary types of scattering this must always be approximated by a finite summation. This means that in any radiative transfer model some type of angular discretization must be applied. For this, one can take two different approaches.

The first is to choose a set of representative discrete directions. In this case the angular integral is approximated by a weighted sum of function values at the discrete directions (discrete ordinate method or DOM, cf. section 3.4). The choice of directions can be based on the optimum derived from the Gaussian integration method, or other criteria are used, such as the even distribution over the sphere applied by Den Dulk (1989). In Den Dulk's models the discrete directions (92 in total) are at the centres of segments shaped as pentagons and hexagons on the sphere. The solid angles associated with these segments are nearly equal. In this case the weights are equal to the cosines of the associate zenith angles.

The second approach is based on an angular segmentation of the sphere of directions, called tessellation. No specific discrete directions are assigned to



these segments. The integral of the RTE is approximated as a sum over N segments, where for each segment numerical integration is used. The choice of a particular type of tessellation can be based on criteria like equal angular intervals (i.e. in zenith and azimuth), equal solid angles (e.g. Den Dulk's "soccer ball" tessellation), or equal weights.

Application of the equal-weight tessellation has the advantage that in this representation the reflectance of a Lambertian surface is given by a uniform matrix, i.e. a matrix of which all elements are equal. Therefore, this type of tessellation has been selected for the $(N+2)$ -stream model discussed in this chapter.

The tessellation which results into equal weights of the segments can be derived from the bidirectional reflectance equation, i.e. Eq. (2.5) of section 2.3:

$$dL_o(\mu_o, \varphi_o) = \rho'(\mu_i, \varphi_i, \mu_o, \varphi_o) L_i(\mu_i, \varphi_i) \mu_i d\Omega_i \quad ,$$

where ρ' = bidirectional reflectance distribution function (BRDF)

L_i = incident radiance from direction (μ_i, φ_i)

$d\Omega_i$ = solid angle of incident flux

dL_o = surface radiance in direction (μ_o, φ_o)

μ_i = $|\cos \theta_i|$, where θ_i is the zenith angle of incidence

μ_o = $|\cos \theta_o|$, where θ_o is the zenith angle of exitence

φ_i = incidence azimuth angle

φ_o = exitence azimuth angle

When all direction arguments are omitted and this equation is integrated over all incidence directions, one finds, with $d\Omega_i = d\mu_i d\varphi_i$:

$$L_o = \int_0^{2\pi} \int_0^1 \rho' L_i \mu_i d\mu_i d\varphi_i \quad .$$

The above integral can also be written as a summation over $N/2$ segments:

$$L_o = \sum_{i=1}^{N/2} \int_{\Delta\varphi_i} \int_{\Delta\mu_i} \rho' L_i \mu_i d\mu_i d\varphi_i \quad ,$$

in which it is assumed that each segment is confined to an interval $\Delta\mu_i$ in



zenith direction and an interval $\Delta\varphi_i$ in azimuth direction.

As an approximation one may assume that the incident radiance L_i is constant within segments. Further, a short notation is introduced for integration of an arbitrary quantity q over a segment by writing

$$\int_{\Delta\varphi_i} \int_{\Delta\mu_i} q(\mu_i, \varphi_i) \mu_i d\mu_i d\varphi_i = \langle q \rangle \frac{1}{2} \Delta\mu_i^2 \Delta\varphi_i \quad .$$

Here $\langle q \rangle$ is called the μ -weighted average of q over the segment, and the product $\frac{1}{2}\Delta\mu_i^2\Delta\varphi_i$ is the weight of the segment. The above integral is now written as

$$L_o = \sum_{i=1}^{N/2} L_i \langle \rho' \rangle \frac{1}{2} \Delta\mu_i^2 \Delta\varphi_i \quad .$$

Segments of equal weight result if one takes equal intervals $\Delta\varphi$ of the azimuth angle and equal intervals $\Delta\mu^2$ of the cosine squared of the zenith angle. In Fig. 5.1 an example of this type of tessellation for a quarter of the upper hemisphere is shown for $\Delta\mu^2 = 1/3$ and $\Delta\varphi = \pi/6$. The indices i and k are used to indicate the zenith zone and the azimuth sector, respectively. Fig. 5.2 shows this numbering of indices for the zenith zones and azimuthal sectors separately.

For the tessellation of the sphere in N segments one obtains

$\frac{1}{2}\Delta\mu^2\Delta\varphi = 2\pi/N$, since for the whole sphere $\Delta\mu^2 = 2$ and $\Delta\varphi = 2\pi$. It is now convenient to express the above equation in the incident flux density field, given by the irradiances from each segment, E_i , where

$$E_i = \int_{\Delta\varphi_i} \int_{\Delta\mu_i} L_i \mu_i d\mu_i d\varphi_i = \langle L_i \rangle \frac{1}{2} \Delta\mu_i^2 \Delta\varphi_i \quad ,$$

but since L_i is assumed to be constant within segments, $E_i = L_i \frac{1}{2}\Delta\mu_i^2\Delta\varphi_i$ and the reflectance equation is written as

$$L_o = \sum_{i=1}^{N/2} \langle \rho' \rangle E_i \quad .$$

The surface radiance L_o can also be integrated over segments to yield a reflected flux density field, given by exitances towards segments j , and which are called E_j . Since $E_j = L_o \frac{1}{2}\Delta\mu_o^2\Delta\varphi_o$, one obtains

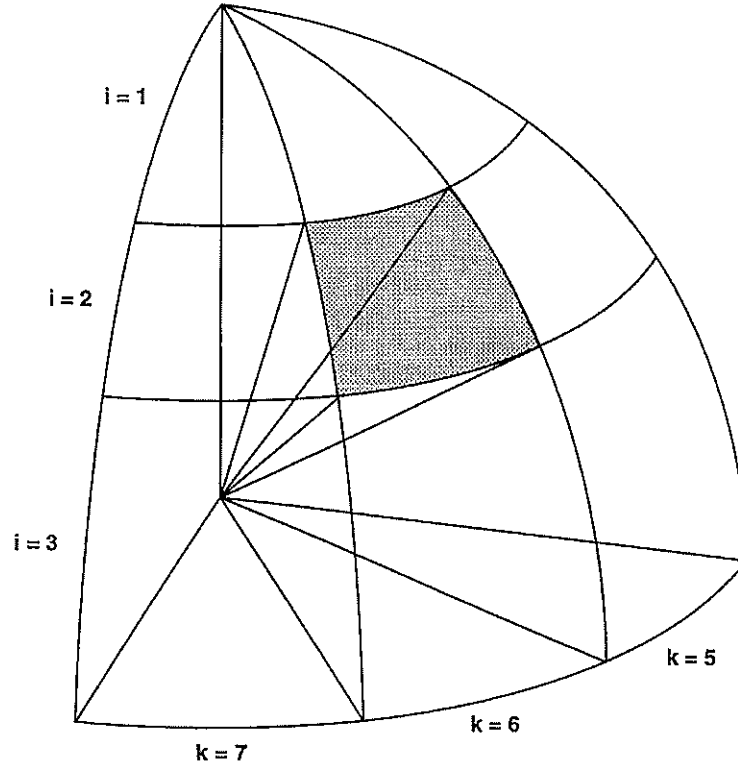


Fig. 5.1 Equal-weight tessellation for a quarter of the upper hemisphere

$$E_j = \sum_{i=1}^{N/2} \langle \langle \rho' \rangle \rangle E_i \frac{1}{2} \Delta \mu_o^2 \Delta \varphi_o = \frac{2\pi}{N} \sum_{i=1}^{N/2} \langle \langle \rho' \rangle \rangle E_i .$$

If the fluxes E_i and E_j are considered elements of vectors E^- and E^+ , respectively, and the surface reflectance is represented by a matrix R , then one can write

$$E^+ = R E^- .$$

The elements of R are denoted as R_{ij} and are given by

$$R_{ij} = \frac{2\pi}{N} \langle \langle \rho'(i,j) \rangle \rangle ,$$

where $\rho'(i,j)$ denotes the bidirectional reflectance distribution function within incident segment i and exitent segment j . For a white Lambertian

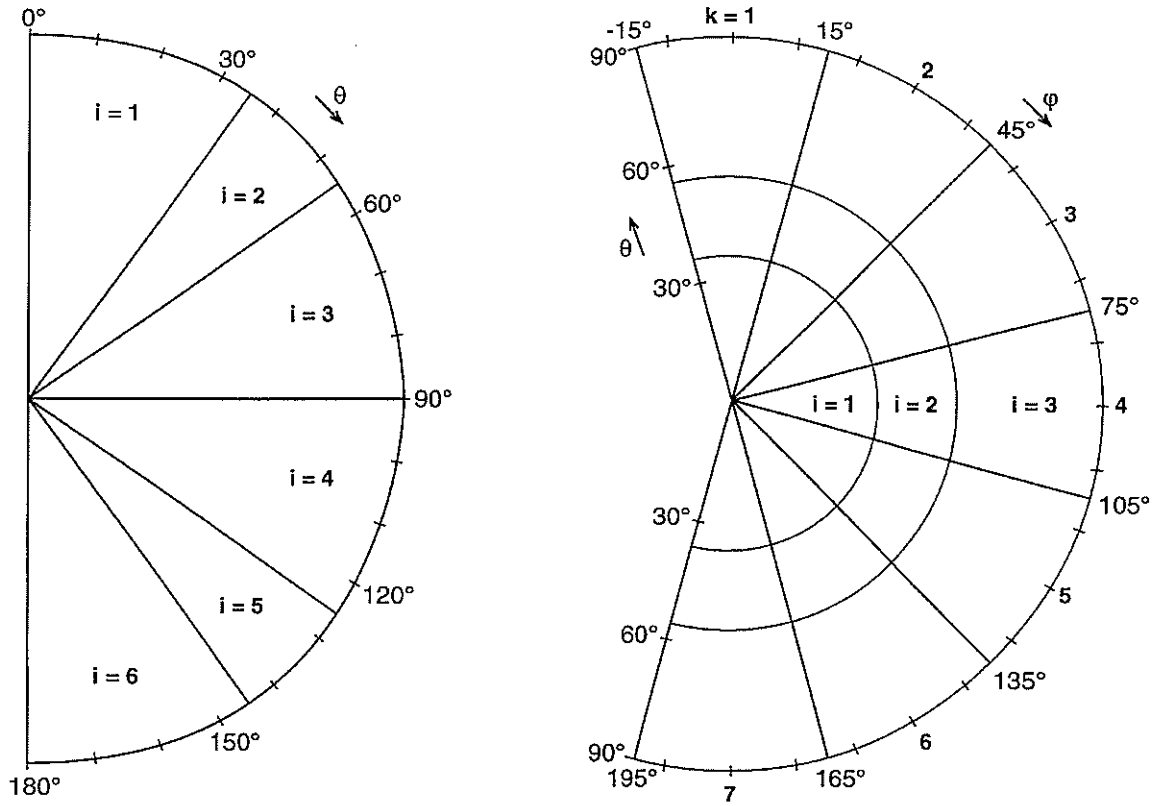


Fig. 5.2 Numbering of zenith and azimuth indices

reflector, ρ' is constant and equal to $1/\pi$ (cf. section 2.3), so in that case all elements of R are equal to $2/N$. The dimension of the vectors E^+ and E^- is $N/2$, and of matrix R it is $N/2 \times N/2$. The total exitance M of the surface is the sum of the elements of vector E^+ . By means of a transposed (row) vector 1^T , of which the elements are all equal to one, this can be written as

$$M = 1^T E^+ = 1^T R E^- .$$

For a white Lambertian reflector this gives

$$M_{WL} = 1^T R_{WL} E^- = (N/2) \cdot (2/N) \cdot 1^T E^- = 1^T E^- ,$$

or the total exitance equals the sum of the incident fluxes, and this is as it should be, because all incident radiation is reflected in this case.



In section 2.3 equations (2.17) and (2.18) were given as the four-stream representation of the reflectance of a surface. These equations can now be generalized to the $(N+2)$ -stream case for the equal-weight tessellation.

In this case the incident fluxes consist of the diffuse flux field from segments i , equal to $E_i = (2\pi/N)L_i$, where L_i is the radiance of the segment, in addition to the specular flux on a horizontal surface, E_s .

The surface radiance L_o then becomes

$$L_o = \sum_{i=1}^{N/2} \langle \rho'(i) \rangle E_i + \rho'(\mu_s, \varphi_s, \mu_o, \varphi_o) E_s, \quad (5.1)$$

where $\rho'(i)$ is the BRDF for incident segment i and exitent direction (μ_o, φ_o) .

Integration over exitent segments j gives

$$E_j = \sum_{i=1}^{N/2} \langle \langle \rho'(i, j) \rangle \rangle E_i (2\pi/N) + \langle \rho'(j) \rangle E_s (2\pi/N). \quad (5.2)$$

It is now convenient to introduce the quantity called "flux-equivalent radiance", defined by $E_o = L_o \frac{1}{2} \Delta\mu_o^2 \Delta\varphi_o = L_o (2\pi/N)$, i.e. the flux density equivalent to the one obtained for a segment with radiance L_o . Eq. (5.1) is now rewritten as

$$E_o = (2\pi/N) \sum_{i=1}^{N/2} \langle \rho'(i) \rangle E_i + (2\pi/N) \rho'(\mu_s, \varphi_s, \mu_o, \varphi_o) E_s.$$

In matrix-vector notation both equations can be written as

$$E_o = r_{so} E_s + r_{do}^T E^-, \quad (5.3)$$

$$E^+ = r_{sd} E_s + R E^-, \quad (5.4)$$

where r_{so} is the bidirectional reflectance times $2/N$, r_{sd} is the specular-diffuse reflectance vector, r_{do}^T is the transposed diffuse-directional reflectance vector, and R is the diffuse reflectance matrix. Note, that for $N = 2$ the four-stream reflectance equations of section 2.3 are obtained again. Also, the definition of any reflectance quantity of Eqs. (5.3) and (5.4) is basically the same, namely



$$r = \frac{2\pi}{N} \langle \langle \rho' \rangle \rangle ,$$

where in this case $\langle \rangle$ denotes μ -weighted averaging over incident or exitent segments, but the size of these segments can be finite as well as infinitesimal. In this respect, the elements of r_{sd} are found if one lets $\frac{1}{2}\Delta\mu_i^2\Delta\varphi_i \rightarrow 0$, those of r_{do}^T for $\frac{1}{2}\Delta\mu_o^2\Delta\varphi_o \rightarrow 0$ and r_{so} is obtained when both conditions apply.

A similar approach can be taken for the description of the optical behaviour of a plane parallel-layer. In that case also transmittances are considered, and the fundamental quantity τ' , or bidirectional transmittance distribution function (BTDF) is used. The result is given by the following set of equations:

$$E_s(b) = \tau_{ss} E_s(t) , \quad (5.5.a)$$

$$E^-(b) = \tau_{sd} E_s(t) + T E^-(t) + R E^+(b) , \quad (5.5.b)$$

$$E^+(t) = \rho_{sd} E_s(t) + R E^-(t) + T E^+(b) , \quad (5.5.c)$$

$$E_o^+(t) = \rho_{so} E_s(t) + \rho_{do}^T E^-(t) + \tau_{do}^T E^+(b) + \tau_{oo} E_o^+(b) , \quad (5.5.d)$$

$$E_o^-(b) = \tau_{so} E_s(t) + \tau_{do}^T E^-(t) + \rho_{do}^T E^+(b) + \tau_{oo} E_o^-(b) , \quad (5.5.e)$$

in which (b) and (t) refer to the bottom and the top of the layer. These equations form the (N+2)-stream equivalent of the associate four-stream case, Eqs. (2.24) of section 2.5. Here, the elements of T , τ_{sd} , τ_{do}^T and τ_{so} are defined by

$$t = \frac{2\pi}{N} \langle \langle \tau' \rangle \rangle , \quad \text{where } t \text{ is the generic for transmittance.}$$

The quantities τ_{ss} and τ_{oo} are direct transmittances and therefore not related with τ' , which results from diffuse scattering. All other quantities of Eqs. (5.5) are interrelated as all are somehow based on either ρ' or τ' .

As $\rho_{so} = (2\pi/N)\rho'$ and $\tau_{so} = (2\pi/N)\tau'$, these relations can also be based on ρ_{so} and τ_{so} . This gives

$$\tau_{sd}(j) = \langle \tau_{so}(j) \rangle$$

$$\rho_{sd}(j) = \langle \rho_{so}(j) \rangle$$



$$\begin{aligned}
 \rho_{do}(i) &= \langle \rho_{so}(i) \rangle \\
 \tau_{do}(i) &= \langle \tau_{so}(i) \rangle \\
 R(i,j) &= \langle \langle \rho_{so}(j) \rangle (i) \rangle = \langle \rho_{sd}(i) \rangle = \langle \rho_{do}(j) \rangle \\
 T(i,j) &= \langle \langle \tau_{so}(j) \rangle (i) \rangle = \langle \tau_{sd}(i) \rangle = \langle \tau_{do}(j) \rangle ,
 \end{aligned}$$

where $\langle \rangle$ denotes μ -weighted averaging over incident segments (i) or exitent segments (j). The above relations are similar to relations given for the four-stream case in section 2.3 and 2.4 (cf. Eqs. (2.11), (2.14), (2.16), (2.21), (2.22) and (2.23)). In that case the segments are the upper and lower hemisphere, so that $1/2 \Delta \mu^2 \Delta \varphi$ equals π , and the operator $\langle \rangle$ applied to a quantity q will give

$$\langle q \rangle = \int_0^{2\pi} \int_0^1 q \mu \, d\mu \, d\varphi / (1/2 \Delta \mu^2 \Delta \varphi) = \pi^{-1} \int_0^{2\pi} \int_0^1 q \mu \, d\mu \, d\varphi ,$$

which is exactly the form in which all these relations appear.

Note that Eqs. (5.5) hold for a plane-parallel layer which is symmetric in its optical properties with respect to the horizontal plane. In other words, when the layer is turned upside down, the same equations apply. Also note, that in Eqs. (5.5.d) and (5.5.e) the quantities ρ_{do}^T , τ_{do}^T and τ_{oo} are only meant to indicate their function; they are not necessarily equal in both instances.

5.3 (N+2)-stream representation of the radiative transfer equation

In this section radiative transfer in a particular kind of medium is considered, namely a medium which is azimuthally isotropic. In such a medium the interception coefficient β depends on the zenith angle, but is independent of the azimuth. Also, the scattering function depends only on the absolute azimuth difference, $|\varphi_o - \varphi_i|$, of the incidence and exitence directions, as well as on the respective zenith angles. Atmospheric layers and homogeneous vegetation canopies such as considered in the SAIL model (cf. chapter 7) are both examples of azimuthally isotropic media. For this type of medium radiative transfer within a plane parallel layer is described by



$$dE_s/dz = -k(\mu_s)E_s, \quad (5.6.a)$$

$$dL_o/dz = -k(\mu_o)L_o + \frac{\gamma(\mu_s, \mu_o, \psi)}{\pi} E_s + \frac{1}{\pi} \int_{4\pi} \gamma(\mu_i, \mu_o, \varphi_o - \varphi_i) L_i \mu_i d\Omega_i, \quad (5.6.b)$$

where E_s = specular irradiance from direction (μ_s, φ_s)

L_i = diffuse incident radiance from direction (μ_i, φ_i)

L_o = exitent radiance in direction (μ_o, φ_o)

k = extinction coefficient

γ = bidirectional scattering coefficient (or area scattering function)

z = metrical depth

$\mu_s = \cos \theta_s$, where θ_s is the zenith angle of specular incidence

$\mu_o = |\cos \theta_o|$, where θ_o is the zenith angle of exitence

$\mu_i = |\cos \theta_i|$, where θ_i is the zenith angle of diffuse incidence

φ_i = azimuth angle of diffuse incidence

φ_o = azimuth angle of exitence

$\psi = |\varphi_o - \varphi_s|$, where φ_s is the azimuth of specular incidence

Strictly, Eq. (5.6.b) applies only to downward exitent radiance L_o , as for upward L_o the signs would have to be reversed for the right-hand side.

With the equal-weight tessellation of the previous section, Eq. (5.6.b) is rewritten as

$$d\pi L_o/dz = -k(\mu_o)\pi L_o + \gamma E_s + \sum_{i=1}^N \langle \gamma(i) \rangle E_i.$$

In terms of the flux-equivalent radiance $E_o = (2\pi/N)L_o$ one obtains

$$dE_o/dz = -k(\mu_o)E_o + \frac{2}{N}\gamma E_s + \frac{2}{N} \sum_{i=1}^N \langle \gamma(i) \rangle E_i,$$

and integration over exitent segments j gives

$$dE_j/dz = -\langle k(j) \rangle E_j + \frac{2}{N} \langle \gamma(j) \rangle E_s + \frac{2}{N} \sum_{i=1}^N \langle \langle \gamma(i,j) \rangle \rangle E_i.$$

One may now consider the scalar flux densities E_s , E_o^+ , E_o^- and the flux



density vectors E^- and E^+ , where E_o^+ and E_o^- are flux-equivalent radiances in upward and downward direction, and E^- and E^+ are vectors of downward and upward flux density within segments. For these the following set of linear differential equations is obtained:

$$dE_s/dz = -k E_s, \quad (5.7.a)$$

$$dE^-/dz = s' E_s - A E^- + B E^+ \quad (5.7.b)$$

$$dE^+/dz = -s E_s - B E^- + A E^+ \quad (5.7.c)$$

$$dE_o^+/dz = -w E_s - v^T E^- - v'^T E^+ + K E_o^+ \quad (5.7.d)$$

$$dE_o^-/dz = w' E_s + v'^T E^- + v^T E^+ - K E_o^- \quad (5.7.e)$$

This system of equations is the generalization of the four-stream differential equations (cf. Eqs. (4.4) of section 4.1.2) to $N+2$ streams. Here, Eq. (5.7.a) describes the extinction of specular flux, and Eqs. (5.7.d-e) represent the radiative transfer equation for upward and downward (flux-equivalent) radiance. Eqs. (5.7.b-c) are the result of integrating Eqs. (5.7.d-e) over existent segments. They are necessary to estimate the diffuse flux field expressed by the vectors E^- and E^+ . Having determined these, they can be substituted in Eqs. (5.7.d-e) in order to estimate the radiances (times π) E_o^+ or E_o^- . The coefficients of Eqs. (5.7) are described by

- k = extinction coefficient for specular flux E_s
- K = extinction coefficient for flux-equivalent radiances E_o^+ and E_o^-
- w, w' = bidirectional scattering coefficient times $2/N$
- s' = vector of specular-diffuse forward scattering coefficients
- s = vector of specular-diffuse backscattering coefficients
- v'^T = (transposed) vector of diffuse-directional forward scattering coefficients
- v^T = (transposed) vector of diffuse-directional backscattering coefficients
- B = matrix of diffuse backscattering coefficients
- A = matrix of diffuse attenuation and forward scattering coefficients.

One may also define a diagonal matrix of diffuse extinction coefficients, κ , and a matrix of diffuse forward scattering coefficients, F . In that case



$A = \kappa - F$. The elements of κ are found by μ -weighted averaging of the extinction coefficient k over a segment, or $\kappa(i) = \langle k(i) \rangle$.

The elements of s' , s , v'^T , v^T , B and F are all found by μ -weighted averaging of the bidirectional scattering coefficient w or w' (times $2/N$) over incident segments, exitent segments, or both, denoted as $\sigma = \langle \langle w \rangle \rangle$, where σ is the generic for any of these scattering coefficients. Here the $\langle \rangle$ operator may refer to finite as well as infinitesimal segments. In practice the numerical calculation of the coefficients is better executed on the basis of the interception coefficient $\beta(\mu)$ and the volume scattering function $\gamma'(\mu_s, \mu_o, \psi)$ as basic quantities, because both $k(\mu)$ and $w(\mu_s, \mu_o, \psi)$ become infinite for $\mu = 0$ and $\mu_s = 0$ or $\mu_o = 0$, i.e. for horizontal directions. For segments adjacent to the horizontal this would complicate the averaging-operation. In terms of $\beta(\mu)$ the diffuse extinction coefficient κ is given by

$$\kappa = \langle k(\mu) \rangle = \langle \beta(\mu)/\mu \rangle, \quad \text{but since}$$

$$\begin{aligned} \langle \beta(\mu)/\mu \rangle &= \int_{\Delta\varphi} \int_{\Delta\mu} \beta(\mu)/\mu \cdot \mu \, d\mu \, d\varphi / (1/2 \Delta\mu^2 \Delta\varphi) \\ &= \frac{\overline{\beta(\mu)} \Delta\mu \Delta\varphi}{1/2 \Delta\mu^2 \Delta\varphi} = \overline{\beta(\mu)} \frac{\Delta\mu}{1/2 \Delta\mu^2}, \end{aligned}$$

$$\kappa = \overline{\beta(\mu)} \cdot 2 \Delta\mu / \Delta\mu^2, \quad (5.8.a)$$

where $\overline{\beta(\mu)}$ is the (normal) average of $\beta(\mu)$ over a segment.

$$\text{Similarly, since } w(\mu_i, \mu_o, \psi) = \frac{2}{N} \gamma(\mu_i, \mu_o, \psi) = \frac{2\pi}{N} \frac{\gamma'(\mu_i, \mu_o, \psi)}{\mu_i \mu_o},$$

scattering coefficients σ are in general found by

$$\begin{aligned} \sigma = \langle \langle w(\mu_i, \mu_o, \psi) \rangle \rangle &= \frac{2\pi}{N} \frac{\int_{\Delta\varphi_o} \int_{\Delta\mu_o} \int_{\Delta\varphi_i} \int_{\Delta\mu_i} \frac{\gamma'(\mu_i, \mu_o, \psi)}{\mu_i \mu_o} \mu_i \, d\mu_i \, d\varphi_i \, \mu_o \, d\mu_o \, d\varphi_o}{1/2 \Delta\mu_i^2 \Delta\varphi_i \, 1/2 \Delta\mu_o^2 \Delta\varphi_o} \\ &= \frac{2\pi}{N} \frac{\overline{\overline{\gamma'(\mu_i, \mu_o, \psi)}}}{\overline{\overline{\gamma'(\mu_i, \mu_o, \psi)}}} \frac{\Delta\mu_i \Delta\varphi_i \Delta\mu_o \Delta\varphi_o}{1/2 \Delta\mu_i^2 \Delta\varphi_i \, 1/2 \Delta\mu_o^2 \Delta\varphi_o}, \quad \text{or} \end{aligned}$$

$$\sigma = \frac{2\pi}{N} \frac{\overline{\overline{\gamma'(\mu_i, \mu_o, \psi)}}}{\overline{\overline{\gamma'(\mu_i, \mu_o, \psi)}}} \frac{2 \Delta\mu_i}{\Delta\mu_i^2} \frac{2 \Delta\mu_o}{\Delta\mu_o^2}. \quad (5.8.b)$$



Here $\overline{\overline{q}}$ denotes (normal) double averaging of quantity q over incident and exitent segments, where these segments can be finite or infinitesimal.

For atmospheres as well as vegetation canopies the functions $\beta(\mu)$ and $\gamma'(\mu_i, \mu_o, \psi)$ are finite for any values of their arguments, so that averaging them poses no special problems.

In particular, for atmospheres $\beta(\mu) = \beta$ is constant and $\gamma'(\mu_i, \mu_o, \psi)$ is given by (cf. section 2.6, Eq. (2.38))

$$\gamma'(\mu_i, \mu_o, \psi) = \frac{\omega\beta}{4\pi} p(\mu_i, \mu_o, \psi) \quad ,$$

where ω is the single scattering albedo and $p(\mu_i, \mu_o, \psi)$ is the scattering phase function. Actually in this case p is only a function of the scattering angle δ , which is the angle between the incidence and exitence direction, and where $\delta = 0$ refers to exact forward scattering and $\delta = \pi$ to exact backscattering. The scattering angle δ can be found from

$$\cos(\pi - \delta) = \cos\theta_i \cos\theta_o + \sin\theta_i \sin\theta_o \cos\psi \quad ,$$

where θ_i and θ_o are the zenith angles associated with μ_i and μ_o . This gives $\delta = 0$ for $\theta_o = \pi - \theta_i$, $\psi = \pi$ and $\delta = \pi$ for $\theta_o = \theta_i$, $\psi = 0$. Now, for atmospheres

$$\kappa = \beta \frac{2\Delta\mu}{\Delta\mu^2} \quad , \text{ and } \sigma = \frac{\omega\beta}{2N} \overline{\overline{p(\delta)}} \frac{2\Delta\mu_i}{\Delta\mu_i^2} \frac{2\Delta\mu_o}{\Delta\mu_o^2} \quad .$$

For vegetation canopies Eqs. (5.8) should be used, since for these $\beta(\mu)$ is not constant in most cases (only for a spherical leaf angle distribution it is).

Integral properties of the coefficients of Eqs. (5.7) are the following:

$$1^T S' + 1^T S = \omega k \quad , \tag{5.9}$$

$$\nu^T 1 + \nu'^T 1 = \omega K \quad , \tag{5.10}$$

$$F 1 + B 1 = \omega \kappa 1 \quad , \tag{5.11}$$



where $\mathbf{1}$ is a vector of all ones, and ω is the single scattering albedo.

These properties result from the law of energy conservation. The above equations express that the total flux scattered in all directions equals a fraction ω of the intercepted flux. For leaf canopies ω is the sum of leaf reflectance and transmittance, $\rho + \tau$. The highest values are found in the near infrared, and here ω is about 0.96. For atmospheres ω can be virtually equal to one, especially in the visible region of the spectrum (conservative scattering). The numerical calculation of the coefficients should be checked by means of Eqs. (5.9) to (5.11). This is necessary as otherwise a "numerical" loss of energy might be introduced.

As a result of Helmholtz's law of reciprocity for scattering, the vector pairs (s', v') and (s, v) interchange position (commute) when the directions of specular incidence and observation are interchanged. In addition, the matrices F and B are symmetric for this reason. The latter property may also be used to check the numerical calculation of coefficients.

For atmospheres as well as leaf canopies all coefficients of Eqs. (5.7) are proportional to some measure of (one-dimensional) density, in units of m^{-1} . For atmospheres this is the interception coefficient β , and for leaf canopies it is the leaf area density, L' , in m^2/m^3 (total one-sided leaf area per unit volume).

Vertical variation of the density is very likely to be present in both types of medium, but this is no problem as long as the other properties show no significant vertical variation. If they do, then a vertical stratification of the medium into (assumed) homogeneous layers is necessary. If not, then one usually assigns an optical thickness, b , to the whole layer, and assumes that the optical density β is constant over the layer, with $\beta = b/h$, where h is the layer height. For leaf canopies one can take the total leaf area index L of the layer, and put $L' = L/h$.

The radiative transfer equation for a homogeneous layer can now be written as

$$\frac{d}{b dx} E = M E \quad , \quad (5.12)$$

where b is generic for optical thickness in the case of the atmosphere and for



the leaf area index in the case of a leaf canopy. Vector E represents all the fluxes of Eqs. (5.7) and matrix M all scattering and extinction coefficients, which in this form are only direction-dependent. The scalar variable x is called the relative optical height, and runs from -1 at the bottom of the layer to zero at the top. As the positive direction for x is upward, Eqs. (5.7) are rewritten as

$$\frac{d}{b dx} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o^+ \\ E_o^- \end{bmatrix} = \begin{bmatrix} k & & & & \\ -s' & A & -B & & \\ s & B & -A & & \\ w & \mathbf{v}^T & \mathbf{v}'^T & -K & \\ -w' & -\mathbf{v}'^T & -\mathbf{v}^T & & K \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o^+ \\ E_o^- \end{bmatrix} \quad (5.13)$$

in which the vacant places of the matrix are zero.

The analytical solution of this system of differential equations, in terms of the reflectance and transmittance quantities defined by Eqs. (5.5), is presented in Appendix A.

5.4 Properties of the analytical solution

In the previous section it has been explained how the equal-weight tessellation of the sphere of directions in N_z zenith zones and N_A azimuth sectors leads to a linear system of differential equations describing the interactions amongst N diffuse fluxes, where $N = N_z N_A$, plus the direct solar flux and a flux in the direction of observation, so $N+2$ "streams" in total.

In Appendix A a solution is presented for the example $N_z = 6$ and $N_A = 12$, but the formulation is general enough to allow an easy extension to arbitrary N_z and N_A , provided both are even. The method is based on eigenvector decomposition. For Eq. (5.12) of the previous section it can be explained by means of the equivalent equation



$$\frac{d}{b dx} Y^{-1} E = Y^{-1} M E = Y^{-1} M Y Y^{-1} E = \Lambda Y^{-1} E \quad ,$$

where Y is the eigenvector matrix and Λ is a diagonal matrix of eigenvalues. The components of $Y^{-1}E$ are linearly weighted combinations of the original fluxes, with the special property that for these flux patterns radiative transfer is simplified to a pure exponential extinction, as if scattering were absent. Apart from their strength, these patterns persist all through the medium, and the relations between them at the top and the bottom of a layer are given by a simple exponential factor. For the determination of all reflectance and transmittance quantities of a layer it is first assumed that all incident fluxes, i.e. the downward fluxes at the top and the upward fluxes at the bottom, are known. From these, and the radiative transfer equation, all exitent fluxes, i.e. the upward fluxes at the top and the downward fluxes at the bottom, are to be calculated. Basically, this is achieved by transforming all fluxes (known and unknown) to eigenvector components, establishing the relations between these at the top and the bottom of the layer (which are simple, see above), and finally transforming back to original fluxes, thus also giving the unknown fluxes.

The eigenvectors and eigenvalues depend on the coefficients of the matrix in Eq. (5.13). As explained in the previous section, these coefficients are independent of the optical density and they depend only on the (bi)directional extinction and scattering properties of the medium. This means that this information is fully represented in the eigenvectors and eigenvalues also, which therefore can be considered constants. The advantage is that, once the eigenvectors and eigenvalues have been determined, one can calculate reflectance and transmittance for various optical thickness values with relatively little effort.

As shown in Appendix A, for the solution of Eq. (5.13) it is not necessary to determine the eigenvectors and eigenvalues of the full matrix, but this can be limited to the matrix which describes only the interactions amongst the diffuse flux vectors, E^- and E^+ , i.e.

$$\frac{d}{b dx} \begin{bmatrix} E^- \\ E^+ \end{bmatrix} = \begin{bmatrix} A & -B \\ B & -A \end{bmatrix} \begin{bmatrix} E^- \\ E^+ \end{bmatrix} \quad (5.14)$$

The dimension of this matrix is $N_Z N_A \times N_Z N_A$, and computation of the eigenvectors and eigenvalues (also called diagonalization) of such a large (asymmetric) matrix can be a problem, as for an $N \times N$ matrix the required processing effort is proportional to N^3 . However, due to three symmetry relations, the number of independent elements of the matrix is only $\frac{1}{2}N_Z (\frac{1}{2}N_Z + 1) (\frac{1}{2}N_A + 1)$ instead of $(N_Z N_A)^2$. For $N_Z = 6$ and $N_A = 12$ this means only 84 independent elements instead of 5184 !

These three symmetry relations are the following:

- I Mirror symmetry with respect to horizontal plane
- II Azimuthal symmetry
- III The reciprocity relation

Symmetry I leaves the radiative transfer equation invariant when the layer is turned upside down. This symmetry is already included in the matrix by the identical **A** and **B** submatrices.

Symmetry II is caused by the choice for an azimuthally isotropic medium. Note that this only means a medium with random azimuthal orientation of the particles, not that scattering is isotropic in azimuth. However, azimuthal scattering is considerably simplified compared to azimuthally non-isotropic media, for which the scattering coefficient depends on the combination of incident and exitent azimuth directions. For azimuthally isotropic media the scattering coefficient depends only on the absolute difference in azimuth of both directions. In the case of N_A azimuth sectors, there are only $\frac{1}{2}N_A + 1$ different absolute differences in azimuth. So, instead of N_A^2 combinations of incident and exitent azimuth sector, one only has to consider $\frac{1}{2}N_A + 1$ absolute differences.

Symmetry III gives equal scattering coefficients when the directions of incidence and exitence are interchanged. Although the interchange of azimuths already causes no effect due to symmetry II, this symmetry still has the additional effect that the associate zenith directions can be interchanged without changing the scattering coefficient. The result is that **A** and **B** are



symmetric matrices.

In Appendix A it is shown that symmetry II can be exploited to achieve a considerable reduction of computational complexity by means of a discrete cosine transform (DCT), which is applied to the N_A azimuthal fluxes per zenith zone. When the sun is supposed to be located in the vertical plane through the centre of an azimuth sector (say, sector one), the diffuse flux field is symmetric with respect to this plane (the principal plane), so that only $\frac{1}{2}N_A + 1$ azimuth sectors have to be considered for a complete description of the diffuse flux field. In this case the DCT is carried out by means of a matrix W of dimension $(\frac{1}{2}N_A + 1) \times (\frac{1}{2}N_A + 1)$. The DCT is a fixed transform, i.e. its coefficients are known *a priori*. With respect to azimuthal scattering, the DCT also acts as an eigenvector matrix, and the associate eigenvalues can easily be found by applying matrix W to the $\frac{1}{2}N_A + 1$ azimuthal scattering coefficients of a zenith zone. By means of the DCT the azimuthal dependence of scattering can be completely separated from the zenithal dependence, giving $\frac{1}{2}N_A + 1$ systems of differential equations of the same form as Eq. (5.14), but for which the dimension of the A and B matrices is only $\frac{1}{2}N_Z \times \frac{1}{2}N_Z$. These matrices now contain the eigenvalues of the DCT-transformed azimuthal scattering coefficients for each combination of incident and existent zenith zone, and are still symmetric due to symmetry III (the reciprocity relation). The DCT-transformed A and B matrices therefore both have $\frac{1}{2} \times \frac{1}{2}N_Z (\frac{1}{2}N_Z + 1)$ independent elements, which leads to $\frac{1}{2}N_Z (\frac{1}{2}N_Z + 1) (\frac{1}{2}N_A + 1)$ elements in total, the same as before the DCT, so there is no loss of information, while storage space is greatly reduced. Also, the eigenvector decomposition is simplified considerably. The computational effort after the DCT is proportional to $(\frac{1}{2}N_A + 1) N_Z^3$ instead of $(N_Z N_A)^3$, which for $N_A = 12$ means a reduction factor of about 250.

The special structure of the matrix of Eq. (5.14), which is asymmetric, can be exploited by deriving the eigenvectors and eigenvalues from the sum $A + B$ and the difference $A - B$, which are symmetric. As shown in Appendix A, first matrix $A + B$ is diagonalized, giving the eigenvector matrix U and the (diagonal) matrix of eigenvalues Δ , such that $U \Delta U^T = A + B$. It also turns out that the eigenvectors and eigenvalues sought are those of the product $(A + B)(A - B)$, which is asymmetric. However, the product-matrix can be transformed into a symmetric matrix by making use of U and



$\Delta^{1/2}$, the matrix which contains the square roots of the diagonal elements of Δ . The transformed matrix equals $\Delta^{1/2}U^T(A-B)U\Delta^{1/2} = V\Lambda^2V^T$, where V is the eigenvector-matrix and Λ^2 the diagonal matrix of eigenvalues. The eigenvalues sought can now be found from the square roots of the diagonal elements of Λ^2 . The eigenvector matrix is given by $Y = U\Delta^{1/2}V$.

Summarized, the above procedure replaces the diagonalization of the asymmetric matrix of size $N_z \times N_z$ by two diagonalizations of symmetric matrices of size $1/2 N_z \times 1/2 N_z$. In terms of computational effort this gives a reduction from N_z^3 to $2 \times (1/2 N_z)^3$, which is a factor of four, but the main advantage is that the involved matrices are symmetric, so that the diagonalizations can be carried out with standard software methods, such as Jacobi's diagonalization of symmetric matrices. In total, a reduction factor of about one thousand in computation time (250×4) has been achieved by the above procedures.

In Jacobi's method the eigenvectors are usually normalized so that they form an orthonormal set, with $UU^T = I$, i.e. the eigenvectors are orthogonal to one another and have unit length. This also means that matrix U has only $1 + \dots + (1/2 N_z - 1)$ independent elements, which equals $1/2 \times 1/2 N_z (1/2 N_z - 1)$ elements. The same holds for matrix V . Matrix $\Delta^{1/2}$ has $1/2 N_z$ diagonal elements, so that the information in matrix Y is represented by $1/2 N_z (1/2 N_z - 1) + 1/2 N_z = (1/2 N_z)^2$ elements. Together with the $1/2 N_z$ diagonal elements of Λ^2 , this gives $1/2 N_z (1/2 N_z + 1)$ elements, which is equal to the number of independent elements in matrices A and B together.

The solution of Eq. (5.14) can be expressed by

$$E^- + E^+ = Y (e^{\Lambda bx} \delta_1 + e^{-\Lambda bx} \delta_2) \quad , \text{ and} \quad (5.15.a)$$

$$E^- - E^+ = Y^{-T} \Lambda (e^{\Lambda bx} \delta_1 - e^{-\Lambda bx} \delta_2) \quad , \quad (5.15.b)$$

where $e^{\Lambda bx}$ is a symbolic notation for the diagonal matrix of the exponential functions $\exp(\lambda_1 bx)$, $\dots$, $\exp(\lambda_n bx)$, with $n = 1/2 N_z$, $Y^{-T} = (Y^{-1})^T = (Y^T)^{-1}$ and δ_1 and δ_2 are constant vectors which have to be determined from the boundary equations. Note that Eqs. (5.15) hold for all $(1/2 N_z + 1)$ DCT-transformed azimuthal "modes" m . Also note, that matrix inversion of Y in order to derive Y^{-T} is not necessary, since it can be found more directly from its constituents by $Y^{-T} = U\Delta^{-1/2}V$.

With the relationships $YY^T = A + B$ and $Y^{-T}\Lambda^2Y^{-1} = A - B$ it can be shown that (cf. Appendix A) Eqs. (5.15) satisfy the differential equation, Eq. (5.14). However, as shown in Appendix A, the above solution does not



correctly represent the case of conservative scattering ($\omega = 1$) for azimuthal mode $m = 0$. Also, the arguments of $e^{-\Lambda b x}$ are positive for the bottom of the layer ($x = -1$), and for large optical thickness b this may give rise to numerical overflow. Therefore, an alternative solution, similar to one proposed by Nakajima & King (1992), is presented. This solution is given by

$$E^- + E^+ = Y [C(x)\delta_1 + S(x)\delta_2] \quad , \quad \text{and} \quad (5.16.a)$$

$$E^- - E^+ = Y^{-T} [\Lambda^2 S(x)\delta_1 + C(x)\delta_2] \quad . \quad (5.16.b)$$

Here, the functions $C(x)$ and $S(x)$ are diagonal matrix functions related with the hyperbolic cosine and sine, respectively. They are defined by

$$C(x) = e^{\Lambda b x} + e^{-\Lambda b(1+x)} \quad \text{and} \quad S(x) = \Lambda^{-1} [e^{\Lambda b x} - e^{-\Lambda b(1+x)}] \quad .$$

Both functions now have only negative or zero arguments in the exponentials. Nakajima & King (1992) apply the discrete ordinates methods (DOM) and define $S(x)$ by $2 S(x) = e^{\Lambda b x} - e^{-\Lambda b(1+x)}$ (in the notation of this thesis). However, the premultiplication by Λ^{-1} is useful in order to express what happens if scattering is conservative. In that case the single scattering albedo $\omega = 1$ and for $m = 0$ this leads to an eigenvalue λ equal to zero. The associate element of $S(x)$ becomes equal to $b(1+2x)$, which is obtained by taking the limit for $\lambda \rightarrow 0$. This means that for $\omega = 1$ and $m = 0$ one of the exponential relations is replaced by a linear one in b and x , which is exactly what is needed in the particular solution obtained for this case. By concentrating the case $\lambda = 0$ in the function $S(x)$ one avoids numerical problems with terms containing λ^{-1} or $\lambda^{-1/2}$, which were present in a previous version of the formulation. By means of the function $S(x)$ the problem of conservative scattering is very well localized, and a single test for $\lambda < \epsilon$, where ϵ is a small number, say 10^{-6} , is sufficient in the computer code to choose between the original function or its limit for $\lambda \rightarrow 0$.

A similar problem as with $\lambda = 0$ can occur after incorporation of the direct solar flux E_s . The extinction coefficient k for this flux gives the exponential decay described by $E_s(x) = E_s(0) e^{k b x}$, where $E_s(0)$ is the incident solar flux at the top of the layer ($x = 0$). In Appendix A it is shown that the case $k = \lambda$, where λ is any of the $\frac{1}{2}N_z(\frac{1}{2}N_A + 1)$ eigenvalues, leads to infinite vectors c and d . However, by replacing these vectors with the vectors p_s and q_s , which are always finite, and introduction of the diagonal matrix function $J(x) = (k - \Lambda)^{-1} [e^{\Lambda b x} - e^{k b x}]$, where $k = kI$, this problem is avoided,



since for $k = \lambda$ the associate element of $J(x)$ can be replaced by the limit for $k \rightarrow \lambda$, which equals $-bx e^{kx}$.

The resultant solution for each mode m is given by

$$\begin{aligned} Y^{-1}(E^{-} + E^{+}) &= C(x)\delta_1 + S(x)\delta_2 + J(x)[\Lambda p_s + q_s]E_s(0) - p_s E_s(0)e^{kx}, \\ Y^T(E^{-} - E^{+}) &= \Lambda^2 S(x)\delta_1 + C(x)\delta_2 + \Lambda J(x)[\Lambda p_s + q_s]E_s(0) - q_s E_s(0)e^{kx}, \end{aligned} \quad (5.17.a-b)$$

where $p_s = (k + \Lambda)^{-1} Y^{-1}(s' - s)$ and $q_s = (k + \Lambda)^{-1} Y^T(s' + s)$.

Here s' and s are the DCT-transformed vectors of scattering coefficients for specular incidence.

To the author's knowledge, the introduction of $J(x)$, p_s and q_s in the solution of the diffuse flux vectors is new and at least not included in the articles of Stamnes *et al.* (1988) or Nakajima & King (1992) on the theory of the discrete ordinates method, which still display factors of $(k - \lambda)^{-1}$ or $(k^2 - \lambda^2)^{-1}$ in certain terms, without addressing possible numerical problems associated with the case $k = \lambda$.

Substitution of $x = 0$ (top of the layer) and $x = -1$ (bottom) in Eqs. (5.17) gives four boundary equations in four unknown vectors, namely $E^{-}(-1)$, $E^{+}(0)$ and the constants δ_1 and δ_2 . Since in Eqs. (5.17) the sum and difference of E^{-} and E^{+} are described, it turns out to be convenient when Eq. (5.5), which defines the reflectances and transmittances of a layer, is also expressed in sums and differences of the diffuse flux vectors. This gives



$$\begin{aligned}
E^-(-1) + E^+(0) &= (\tau_{sd} + \rho_{sd})E_s(0) + (T + R)[E^-(0) + E^+(-1)] \quad ; \\
E^-(-1) - E^+(0) &= (\tau_{sd} - \rho_{sd})E_s(0) + (T - R)[E^-(0) - E^+(-1)] \quad ; \\
E_o^+(0) + E_o^-(-1) &= (\rho_{so} + \tau_{so})E_s(0) + (\rho_{do}^T + \tau_{do}^T)[E^-(0) + E^+(-1)] \\
&\quad + \tau_{oo}[E_o^+(-1) + E_o^-(0)] \quad ; \\
E_o^+(0) - E_o^-(-1) &= (\rho_{so} - \tau_{so})E_s(0) + (\rho_{do}^T - \tau_{do}^T)[E^-(0) - E^+(-1)] \\
&\quad + \tau_{oo}[E_o^+(-1) - E_o^-(0)] \quad .
\end{aligned}$$

As shown in Appendix A, the boundary equations can be brought into the same form as the former two of the above equations, thus giving solutions for $T + R$, $T - R$, $\tau_{sd} + \rho_{sd}$ and $\tau_{sd} - \rho_{sd}$. Two methods can be used for this. In the first method δ_1 and δ_2 are eliminated from the boundary equations, whereas in the second δ_1 and δ_2 are first solved (expressed in known fluxes $E^-(0)$, $E^+(-1)$ and $E_s(0)$) and then substituted in Eqs. (5.17) in order to derive the sum and difference of $E^-(-1)$ and $E^+(0)$.

Surprisingly, both methods lead to apparently different answers. However, in Appendix A it is shown that the apparent differences are caused by the symmetry of the matrices T and R , and that the expressions found for both methods are equivalent. Symmetry of T and R is also required in order to fulfil the reciprocity relations for the diffuse fluxes.

The quantities $\rho_{so} + \tau_{so}$, $\rho_{so} - \tau_{so}$, $\rho_{do}^T + \tau_{do}^T$ and $\rho_{do}^T - \tau_{do}^T$ can be found by solving $E_o^+(0)$ and $E_o^-(-1)$ and expressing the results in the incident fluxes $E_s(0)$, $E^-(0)$, $E^+(-1)$, $E_o^+(-1)$ and $E_o^-(0)$. For this, it is assumed that E_o^+ and E_o^- are in exactly opposite directions, so that the following differential equations apply:

$$\begin{aligned}
\frac{d}{b dx} E_o^+ &= w E_s + v^T E^- + v'^T E^+ - K E_o^+ \quad , \text{ and} \\
\frac{d}{b dx} E_o^- &= -w' E_s - v'^T E^- - v^T E^+ + K E_o^- \quad .
\end{aligned}$$

Analytical solutions of these equations are given by



$$E_o^+(0) = \int_{-1}^0 (wE_s + v^T E^- + v'^T E^+) e^{Kbx} dx + E_o^+(-1)e^{-Kb} \quad , \quad \text{and}$$

$$E_o^-(-1) = \int_{-1}^0 (w'E_s + v'^T E^- + v^T E^+) e^{-Kb(1+x)} dx + E_o^-(0)e^{-Kb} \quad .$$

From these, one immediately finds $\tau_{oo} = e^{-Kb}$. The other quantities follow from both integrals, in which the analytical expressions for the diffuse flux vectors E^- and E^+ as obtained from Eqs. (5.17) must be substituted. For this, the constant vectors δ_1 and δ_2 must also be known (expressed in known incident fluxes); hence the need for the second method of expressing the boundary equations. Once the above equations have been expressed in $E_s(0)$, $E^-(0) + E^+(-1)$ and $E^-(0) - E^+(-1)$, the quantities $\rho_{so} + \tau_{so}$, $\rho_{so} - \tau_{so}$, $\rho_{do}^T + \tau_{do}^T$ and $\rho_{do}^T - \tau_{do}^T$ can be determined.

Much effort has been spent in finding expressions which are numerically "safe" and in which the reciprocity relations are automatically fulfilled. In Appendix A the derivations are presented in detail. The final results can be summarized as follows:

I Dependence on optical thickness b

The following diagonal matrix functions depend on b :

$$\begin{aligned} C &= I + e^{-\Lambda b} \quad ; \quad S = \Lambda^{-1} [I - e^{-\Lambda b}] \quad ; \\ J_1(k) &= (k - \Lambda)^{-1} [e^{-\Lambda b} - e^{-kb}] \quad ; \quad J_1(K) = (K - \Lambda)^{-1} [e^{-\Lambda b} - e^{-Kb}] \quad ; \\ D^- &= (K - k)^{-1} [J_1(k) - J_1(K)] \quad . \end{aligned}$$

The following scalar functions depend on b :

$$\begin{aligned} e^{-kb} \quad ; \quad e^{-Kb} \quad ; \\ Q^+ &= (K + k)^{-1} [1 - e^{-(K+k)b}] \quad ; \quad Q^- = (K - k)^{-1} [e^{-kb} - e^{-Kb}] \quad . \end{aligned}$$

The above diagonal matrices and scalars are all that is necessary to describe the dependence on optical thickness.



II Direction-dependent vectors

a) The following vectors depend on the solar zenith angle:

$$\mathbf{p}_s = (\mathbf{k} + \Lambda)^{-1} \mathbf{Y}^{-1} (\mathbf{s}' - \mathbf{s}) \quad ; \quad \mathbf{q}_s = (\mathbf{k} + \Lambda)^{-1} \mathbf{Y}^T (\mathbf{s}' + \mathbf{s}) \quad ;$$

$$\mathbf{g}_1^+ = (1 + e^{-kb}) \mathbf{p}_s - \mathbf{J}_1(k) (\Lambda \mathbf{p}_s + \mathbf{q}_s) \quad ;$$

$$\mathbf{g}_2^+ = (1 - e^{-kb}) \mathbf{q}_s + \Lambda \mathbf{J}_1(k) (\Lambda \mathbf{p}_s + \mathbf{q}_s) \quad ;$$

$$\mathbf{g}_1^- = (1 - e^{-kb}) \mathbf{p}_s + \mathbf{J}_1(k) (\Lambda \mathbf{p}_s + \mathbf{q}_s) \quad ;$$

$$\mathbf{g}_2^- = (1 + e^{-kb}) \mathbf{q}_s - \Lambda \mathbf{J}_1(k) (\Lambda \mathbf{p}_s + \mathbf{q}_s) \quad .$$

b) The following vectors depend on the viewing direction (zenith angle and relative azimuth)

$$\mathbf{p}_v = (\mathbf{K} + \Lambda)^{-1} \mathbf{Y}^{-1} (\mathbf{v}' - \mathbf{v}) \quad ; \quad \mathbf{q}_v = (\mathbf{K} + \Lambda)^{-1} \mathbf{Y}^T (\mathbf{v}' + \mathbf{v}) \quad ;$$

$$\mathbf{h}_1^+ = (1 + e^{-Kb}) \mathbf{p}_v - \mathbf{J}_1(K) (\Lambda \mathbf{p}_v + \mathbf{q}_v) \quad ;$$

$$\mathbf{h}_2^+ = (1 - e^{-Kb}) \mathbf{q}_v + \Lambda \mathbf{J}_1(K) (\Lambda \mathbf{p}_v + \mathbf{q}_v) \quad ;$$

$$\mathbf{h}_1^- = (1 - e^{-Kb}) \mathbf{p}_v + \mathbf{J}_1(K) (\Lambda \mathbf{p}_v + \mathbf{q}_v) \quad ;$$

$$\mathbf{h}_2^- = (1 + e^{-Kb}) \mathbf{q}_v - \Lambda \mathbf{J}_1(K) (\Lambda \mathbf{p}_v + \mathbf{q}_v) \quad .$$

III Bidirectional quantities

Apart from w and w' , the bidirectional scattering coefficients, the following (scalar) quantities are bidirectional:

$$a^+ = \frac{1}{4} (\mathbf{Y} \mathbf{h}_1^+ + \mathbf{Y}^{-T} \mathbf{h}_2^+)^T (\mathbf{R} + \mathbf{T}) (\mathbf{Y} \mathbf{g}_1^+ + \mathbf{Y}^{-T} \mathbf{g}_2^+) + \frac{1}{4} (\mathbf{h}_2^{+T} \mathbf{Y}^{-1} \mathbf{Y}^{-T} \mathbf{g}_2^+ - \mathbf{h}_1^{+T} \mathbf{Y}^T \mathbf{Y} \mathbf{g}_1^+) \quad ;$$

$$a^- = \frac{1}{4} (\mathbf{Y} \mathbf{h}_1^- + \mathbf{Y}^{-T} \mathbf{h}_2^-)^T (\mathbf{R} - \mathbf{T}) (\mathbf{Y} \mathbf{g}_1^- + \mathbf{Y}^{-T} \mathbf{g}_2^-) + \frac{1}{4} (\mathbf{h}_2^{-T} \mathbf{Y}^{-1} \mathbf{Y}^{-T} \mathbf{g}_2^- - \mathbf{h}_1^{-T} \mathbf{Y}^T \mathbf{Y} \mathbf{g}_1^-) \quad .$$

The final results are given by



$$\mathbf{R} + \mathbf{T} = [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\Lambda^2\mathbf{Y}^{-1}]^{-1} [\mathbf{C}\mathbf{Y}^T - \mathbf{S}\Lambda^2\mathbf{Y}^{-1}] \quad ;$$

$$\mathbf{R} - \mathbf{T} = [\mathbf{S}\mathbf{Y}^T + \mathbf{C}\mathbf{Y}^{-1}]^{-1} [\mathbf{S}\mathbf{Y}^T - \mathbf{C}\mathbf{Y}^{-1}] \quad ;$$

$$\rho_{sd} + \tau_{sd} = [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\Lambda^2\mathbf{Y}^{-1}]^{-1} [\mathbf{C}\mathbf{g}_2^+ - \mathbf{S}\Lambda^2\mathbf{g}_1^+] \quad ;$$

$$\rho_{sd} - \tau_{sd} = [\mathbf{S}\mathbf{Y}^T + \mathbf{C}\mathbf{Y}^{-1}]^{-1} [\mathbf{S}\mathbf{g}_2^- - \mathbf{C}\mathbf{g}_1^-] \quad ;$$

$$\rho_{do} + \tau_{do} = [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\Lambda^2\mathbf{Y}^{-1}]^{-1} [\mathbf{C}\mathbf{h}_2^+ - \mathbf{S}\Lambda^2\mathbf{h}_1^+] \quad ;$$

$$\rho_{do} - \tau_{do} = [\mathbf{S}\mathbf{Y}^T + \mathbf{C}\mathbf{Y}^{-1}]^{-1} [\mathbf{S}\mathbf{h}_2^- - \mathbf{C}\mathbf{h}_1^-] \quad ;$$

$$\begin{aligned} \rho_{so}^d = & \frac{1}{2}(a^+ + a^-) + \frac{1}{4}(\Lambda\mathbf{p}_v - \mathbf{q}_v)^T \mathbf{J}_1(k) e^{-Kb} (\Lambda\mathbf{p}_s + \mathbf{q}_s) \\ & + \frac{1}{4}(\Lambda\mathbf{p}_v + \mathbf{q}_v)^T \mathbf{J}_1(K) e^{-kb} (\Lambda\mathbf{p}_s - \mathbf{q}_s) \\ & + \frac{1}{4}\mathbf{Q}^+ [(K-k)(\mathbf{p}_v^T \mathbf{q}_s - \mathbf{q}_v^T \mathbf{p}_s) - 2\mathbf{p}_v^T \Lambda^2 \mathbf{p}_s + 2\mathbf{q}_v^T \mathbf{q}_s] \quad ; \end{aligned}$$

$$\begin{aligned} \tau_{so}^d = & \frac{1}{2}(a^+ - a^-) + \frac{1}{4}(\Lambda\mathbf{p}_v + \mathbf{q}_v)^T (K+k+2\Lambda) \mathbf{D}^- (\Lambda\mathbf{p}_s + \mathbf{q}_s) \\ & + \frac{1}{4}\mathbf{Q}^- [\mathbf{p}_v^T (K+k+2\Lambda) \mathbf{q}_s + \mathbf{q}_v^T (K+k+2\Lambda) \mathbf{p}_s] \quad . \end{aligned}$$

All of the above quantities apply to all azimuthal modes m .

The remaining quantities are:

$$\tau_{ss} = e^{-kb} \quad ; \quad \tau_{oo} = e^{-Kb} \quad ;$$

$$\rho_{so}^s = w\mathbf{Q}^+ \quad ; \quad \tau_{so}^s = w'\mathbf{Q}^- \quad ;$$

$$\rho_{so} = \rho_{so}^s + \sum_{m=0}^{\frac{1}{2}N_A} \rho_{so}^d(m) \quad ; \quad \tau_{so} = \tau_{so}^s + \sum_{m=0}^{\frac{1}{2}N_A} \tau_{so}^d(m) \quad .$$

Here, ρ_{so}^s and τ_{so}^s are single scattering contributions, ρ_{so}^d and τ_{so}^d multiple scattering contributions to the bidirectional reflectance and transmittance, respectively.

The reciprocity relations are automatically fulfilled, as the matrices $\mathbf{R}$ and $\mathbf{T}$ are symmetric, the couples (ρ_{sd}, ρ_{do}) and (τ_{sd}, τ_{do}) commute when the direc-



tions of view and solar incidence are interchanged, and ρ_{so} and τ_{so} are completely symmetric with respect to both directions.

Potential numerical problems are concentrated in the functions S , $J_1(k)$, $J_1(K)$, Q^- and D^- .

The case of conservative scattering, giving an eigenvalue $\lambda = 0$ for $m = 0$, only affects the diagonal matrix function S . This case is recovered by replacing the associate element of S by its limit for $\lambda \rightarrow 0$, which equals

$$\lim_{\lambda \rightarrow 0} \frac{1 - e^{-\lambda b}}{\lambda} = \frac{1 - (1 - \lambda b)}{\lambda} = b.$$

The other possible cases are $k = \lambda$, $K = \lambda$, $K = k$ and $K = k = \lambda$, where λ is any of the eigenvalues.

When $k = \lambda$, the associate element of $J_1(k)$ is replaced by the limit for $k \rightarrow \lambda$, which equals $b e^{-kb}$.

When $K = \lambda$, the associate element of $J_1(K)$ is replaced by the limit for $K \rightarrow \lambda$, which equals $b e^{-Kb}$.

The above limits must also be substituted in D^- , when applicable.

When $K = k$, the function Q^- is replaced by the limit for $K \rightarrow k$, which equals $b e^{-kb}$. In this case D^- is replaced by an alternative expression, since the given expression becomes indeterminate for $K \rightarrow k$. One can write:

$$\begin{aligned} (K - \Lambda)D^- &= (K - k)^{-1} [(K - k + k - \Lambda)J_1(k) - e^{-\Lambda b} + e^{-Kb}] \\ &= J_1(k) + (K - k)^{-1} [e^{-\Lambda b} - e^{-kb} - e^{-\Lambda b} + e^{-Kb}] = J_1(k) - Q^-, \\ \text{or } D^- &= (K - \Lambda)^{-1} [J_1(k) - Q^-]. \end{aligned}$$

For $K = k$, this gives $D^- = (K - \Lambda)^{-1} [J_1(K) - b e^{-Kb}]$, provided $K \neq \lambda$.



When $K = k = \lambda$, the associate element of D^- is written as

$$\begin{aligned}
 & (K-\lambda)^{-1} \left[\frac{e^{-\lambda b} - e^{-kb}}{k-\lambda} - \frac{e^{-kb} - e^{-Kb}}{K-k} \right] \\
 &= (K-\lambda)^{-1} e^{-kb} \left[\frac{e^{(k-\lambda)b} - 1}{k-\lambda} - \frac{1 - e^{-(K-k)b}}{K-k} \right] \\
 &\approx (K-\lambda)^{-1} e^{-kb} \left[\frac{(k-\lambda)b + \frac{1}{2}(k-\lambda)^2 b^2}{k-\lambda} - \frac{(K-k)b - \frac{1}{2}(K-k)^2 b^2}{K-k} \right] \\
 &= (K-\lambda)^{-1} e^{-kb} [b + \frac{1}{2}(k-\lambda)b^2 - b + \frac{1}{2}(K-k)b^2] = \frac{1}{2} b^2 e^{-kb} .
 \end{aligned}$$

From the above analyses of potential numerical difficulties it is concluded that in the present formulation these are very well localized in a limited number of functions which can be replaced by appropriate limits whenever a special case is encountered. This has been realized by always combining the factors λ^{-1} , $(k-\lambda)^{-1}$, $(K-\lambda)^{-1}$ and $(K-k)^{-1}$ with associate differences of exponential functions, thus guaranteeing that the combinations remain finite when the denominator becomes zero. Previous versions of the theory (Verhoef, 1988) did not have this property and therefore were more sensitive to numerical problems caused by the inaccurate calculation of differences of very large numbers in a computer.

5.5 Mathematical validation

The expressions for $R+T$ and $R-T$ of the previous section have been checked by comparing them with corresponding expressions given by Nakajima & King (1992). Apart from notational differences, they are identical. In their article, these authors study the behaviour of R and T for very large, but not infinite, optical thickness, which leads to asymptotic expressions that are relatively simple.

In this section approximations for small and infinite optical thickness will be given for the quantities $R+T$, $R-T$, $\tau_{sd} + \rho_{sd}$, $\tau_{sd} - \rho_{sd}$, ρ_{so} and τ_{so} .

Here, quantities approximated by expressions for small optical thickness b are denoted with an ϵ -subscript and those for infinite optical thickness by an



∞ -subscript. Derivations are given in Appendix B. The results are the following:

$$(R+T)_\epsilon = I - b(A-B) \quad ; \quad (R-T)_\epsilon = -I + b(A+B) \quad ;$$

$$(\rho_{sd} + \tau_{sd})_\epsilon = b(s' + s) \quad ; \quad (\rho_{sd} - \tau_{sd})_\epsilon = b(s - s') \quad ;$$

$$(\rho_{so}^d)_\epsilon = 1/2 b^2 (v'^T s + v^T s') \quad ; \quad (\rho_{so}^s)_\epsilon = b w \quad ;$$

$$(\tau_{so}^d)_\epsilon = 1/2 b^2 (v'^T s' + v^T s) \quad ; \quad (\tau_{so}^s)_\epsilon = b w' \quad ;$$

$$R_\infty = [Y^T + \Lambda Y^{-1}]^{-1} [Y^T - \Lambda Y^{-1}] \quad ; \quad T_\infty = O \quad ;$$

$$(\rho_{sd})_\infty = [Y^T + \Lambda Y^{-1}]^{-1} [q_s - \Lambda p_s] \quad ; \quad (\tau_{sd})_\infty = 0 \quad ;$$

$$\begin{aligned} (\rho_{so}^d)_\infty &= \frac{1}{4} [Y^{-T} q_v + Y p_v]^T R_\infty [Y^{-T} q_s + Y p_s] \\ &\quad + \frac{1}{4} [q_v^T Y^{-1} Y^{-T} q_s - p_v^T Y^T Y p_s] \\ &\quad + \frac{1}{4} (K+k)^{-1} [(K-k)(p_v^T q_s - q_v^T p_s) - 2 p_v^T \Lambda^2 p_s + 2 q_v^T q_s] \quad ; \\ (\tau_{so}^d)_\infty &= 0 \quad . \end{aligned}$$

These results are interpreted as follows:

For small optical thickness one obtains $R = bB$, $\rho_{sd} = bs$, $\tau_{sd} = bs'$, $\rho_{so} = bw$ and $\tau_{so} = bw'$. In other words, these quantities are all proportional to the associate scattering coefficients (matrix, vector or scalar) and the optical thickness, as could be expected.

For T one finds $T = I - bA$, so for zero optical thickness $T = I$, the identity matrix, indicating total transparency. This means that the direct transmission of diffuse flux (without interaction) is included in matrix T .

Since $A = \kappa - F$, where κ is the diagonal matrix of extinction coefficients, and F the matrix of forward scattering coefficients for diffuse fluxes, the result for small optical thickness can be written as $T = I - b\kappa + bF$, where the part $I - b\kappa$ can be attributed to direct transmittance and the part bF to diffuse forward scattering.

The multiple scattering contributions to ρ_{so} and τ_{so} , denoted by ρ_{so}^d and τ_{so}^d , are proportional to the square of the optical thickness and inner products of vectors of scattering coefficients. These inner products describe the conver-



sion (by double scattering)

direct solar flux $\Rightarrow$ diffuse flux $\Rightarrow$ observed radiance ,

which explains why this contribution depends quadratically on the optical thickness, as a minimum of two scattering events are required for this interaction, and since for small optical thickness the probability of a single scattering event is proportional to b , the probability of two scattering events will be proportional to b^2 . From this it can also be concluded that for small optical thickness the single scattering contribution will dominate the result, as in this case b^2 is much smaller than b . However, for larger b , say $b > 1$, the multiple scattering contribution grows much faster than the single scattering one, so eventually this will become the dominant effect.

For infinite optical thickness all transmittances have reduced to zero, as expected, and the reflectances have reached stable values. These depend on the directional scattering properties of the medium, the absorption characteristics and the directions of incidence and exitence.

5.6 Conclusion

Four-stream radiative transfer theory, which is the basis for the SAIL vegetation canopy reflectance model and for a similar atmospheric scattering model, has been generalized to a so-called $(N+2)$ -stream theory. Starting point for the new model is the tessellation of the sphere of directions in N segments of equal radiometric weight in order to describe the diffuse flux field. The result of this is a system of $N+2$ linear differential equations by which radiative transfer is described. In the analytical solution of this system of differential equations the symmetry relations of bidirectional scattering in azimuthally isotropic media are fully exploited in order to achieve a considerable reduction of computational complexity and required storage space. In the formulation of the theory special attention has been paid to potential numerical problems related with equal or nearly equal eigenvalues. These problems have been identified and solved. Also, the resulting reflectance and transmittance quantities are formulated in such a way, that the reciprocity relations amongst them follow automatically. The model has been validated mathematically by analyzing its behaviour for small and infinite optical



thickness.

Contrary to DOM-based multistream models, which use Gaussian integration and without exception rely on the medium being isotropic, like is the case for atmospheres, the $(N+2)$ -stream model presented in the previous sections can also be applied to non-isotropic media such as (azimuthally isotropic) leaf canopies. The model's output consists of all the necessary reflectance and transmittance quantities (matrices, vectors and scalars) required for application of the adding method. Thus, it will now be possible to build a numerically accurate coupled model of vegetation bidirectional reflectance and atmospheric scattering. However, this task is left for the future, as the model so far has only been implemented for the study of atmospheric scattering. In the next chapter the results of a number of simulations with this model are presented. These simulations were intended to investigate the functioning of the model under a wide variety of input conditions, rather than to provide realistic simulations for the terrestrial atmosphere.

6 SIMULATION RESULTS FOR (N+2)-STREAM ATMOSPHERE MODEL

6.1 Introduction

The theory presented in the previous chapter has been implemented in several computer programs in order to explore its performance for atmospheric layers of different composition and optical thickness. The (N+2)-stream model has been tested with respect to the following aspects:

- Eigenvalues and eigenvectors for aerosol and Rayleigh scattering
- Infinite reflectance matrix
- The effect of the single scattering albedo on hemispherical reflectance, transmittance and layer absorbance as a function of optical thickness and solar zenith angle
- Bidirectional reflectance and transmittance profiles for different solar zenith angles and a wide range of optical thicknesses
- The reciprocity relations
- The contributions due to single and multiple scattering
- Comparison of simulation results for $N = 72$ and $N = 2$

As a prototype for the phase function of aerosol scattering, the one for water haze M (maritime type) has been chosen (Deirmendjian, 1969). For a number of different wavelengths this phase function is shown in chapter 9 (Fig. 9.3), together with the phase function for Rayleigh scattering.

In this chapter the results of model simulations for both types of scattering (aerosol and Rayleigh) will be discussed.

6.2 Eigenvalues and eigenvectors

The eigenvalues and eigenvectors of the matrix of diffuse extinction and scattering depend only on the scattering phase function and the single scattering albedo, as all extinction and scattering coefficients have been defined per unit of optical thickness (cf. chapter 5). For aerosol scattering and Rayleigh scattering the phase functions are clearly different from one another. However, the single scattering albedo for Rayleigh scattering can be taken equal to one, and for aerosol scattering it is usually close to one, so



the main differences in eigenvalues and eigenvectors associated with both types of scattering are expected to be caused by the differences in the respective phase functions.

Numerical results for a few cases are presented in Tables 6.1 to 6.3. In each table the eigenvalues are listed in increasing order for each of the 7 azimuthal modes m (0 to 6, cf. section 5.6 and Appendix A) associated with the discrete cosine transform (DCT). The associate eigenvectors are on the right in transposed form. Table 6.1 gives the results for aerosol Haze M at a wavelength of 700 nm and a single scattering albedo ω_A of one (no absorption). In Table 6.2 a small absorption ($\omega_A = 0.99$) has been introduced in order to show the effect of this. Table 6.3 shows the results for Rayleigh scattering. In this case there is no absorption either but the scattering phase function is highly different. The results can be interpreted as follows.

The smallest eigenvalue in Table 6.1 is 0.000018, where it should be zero for this case of conservative scattering. The reason that the calculation gives a number slightly different from zero is numerical inaccuracy and the fact that single precision arithmetic has been used. The associate eigenvector has equal components, so it is a scalar multiple of the 1-vector, as expected.

From Table 6.2 it is concluded that a small absorption of only one percent has a relatively large effect on the smallest eigenvalue and only a small (relative) effect on the other eigenvalues. The eigenvectors are also slightly affected.

The results for Rayleigh scattering are presented in Table 6.3. Here the smallest eigenvalue is zero indeed and the associate eigenvector is again a multiple of the 1-vector. Why the limited numerical precision appears to have less effect in the case of Rayleigh scattering is not known, but perhaps this is related to the much lower degree of anisotropy of this type of scattering.

A peculiarity with Rayleigh scattering is revealed by the fact that for modes m greater than 2 the set of eigenvalues and eigenvectors becomes independent of m . Also, the matrix of eigenvectors becomes diagonal and the non-zero element of each eigenvector is the square root of the associate eigenvalue. On closer examination, it appears that the eigenvalues for $m > 2$ are equal to the diffuse extinction coefficients, associated with the zenith sectors 1 to 3. These are equal to $6[1 - (2/3)^{1/2}]$, $6[(2/3)^{1/2} - (1/3)^{1/2}]$ and $6(1/3)^{1/2}$, respectively, which follows from Eq. (5.8.a) of section 5.3 if the interception coefficient is set equal to one and N_z equals 3.



Table 6.1 Eigenvalues and eigenvectors for aerosol Haze M at 700 nm and $\omega_A = 1$

Mode	Eigenvalue	Eigenvector (transposed)		
0	0.000018	0.353876	0.353876	0.353876
	0.752800	-0.579229	0.548623	0.569988
	1.452165	0.093491	-0.628659	1.246135
1	0.522447	0.521157	0.440577	0.147862
	0.906949	0.659429	-0.609388	-0.234186
	1.653504	0.008357	-0.327944	1.004333
2	0.796565	0.734090	0.499201	0.155894
	1.027490	-0.577214	0.813171	0.243541
	1.937718	0.000531	-0.327944	1.493460
3	0.947098	0.919480	0.315496	0.038316
	1.135971	-0.349006	0.998677	0.112927
	2.317190	-0.005054	-0.192723	1.433406
4	1.015264	0.994541	0.161107	0.012437
	1.219955	-0.177043	1.088127	0.074136
	2.605434	-0.001918	-0.102168	1.661038
5	1.045875	1.018577	0.090995	0.003841
	1.271274	-0.100424	1.122074	0.042104
	2.791248	-0.000774	-0.065626	1.637211
6	1.054875	1.024494	0.072079	0.002575
	1.288225	-0.079680	1.131623	0.036674
	2.853212	-0.000374	-0.052912	1.715615



Table 6.2 Eigenvalues and eigenvectors for aerosol Haze M at 700 nm and $\omega_A = 0.99$

Mode	Eigenvalue	Eigenvector (transposed)		
0	0.087075	0.366556	0.355929	0.344789
	0.764218	-0.577790	0.563030	0.560036
	1.468339	0.087101	-0.619575	1.258991
1	0.531683	0.530962	0.442335	0.146073
	0.912917	0.654607	-0.621376	-0.232034
	1.669623	0.005945	-0.557834	1.106858
2	0.801145	0.741017	0.494044	0.151131
	1.031399	-0.570306	0.821000	0.239811
	1.952421	-0.000210	-0.322559	1.498396
3	0.949026	0.922309	0.310447	0.037219
	1.138879	-0.343413	1.002296	0.111504
	2.328544	-0.005109	-0.189923	1.438404
4	1.016208	0.995427	0.158570	0.012098
	1.222102	-0.174318	1.089607	0.073117
	2.614001	-0.001924	-0.100860	1.663189
5	1.046452	1.018978	0.089677	0.003748
	1.272914	-0.099004	1.122954	0.041589
	2.797973	-0.000774	-0.064827	1.639610
6	1.055353	1.024799	0.071067	0.002513
	1.289695	-0.078587	1.132362	0.036214
	2.859319	-0.000376	-0.052288	1.717138



Table 6.3 Eigenvalues and eigenvectors for Rayleigh scattering

Mode	Eigenvalue	Eigenvector (transposed)		
0	0.000000	0.726654	0.726654	0.726654
	1.235408	0.732808	-0.837762	-0.283073
	2.375595	-0.189702	-0.452778	1.689954
1	1.090499	1.047658	0.066024	0.016588
	1.404055	-0.057992	1.195849	0.032588
	3.413692	-0.008324	-0.021531	1.860850
2	1.100437	1.049266	0.007944	0.006344
	1.426468	-0.007014	1.197688	0.029389
	3.326881	-0.003466	-0.018942	1.860967
3	1.101021	1.049295	0.000000	0.000000
	1.434878	0.000000	1.197864	0.000000
	3.464102	0.000000	0.000000	1.861210
4	1.101021	1.049295	0.000000	0.000000
	1.434878	0.000000	1.197864	0.000000
	3.464102	0.000000	0.000000	1.861210
5	1.101021	1.049295	0.000000	0.000000
	1.434878	0.000000	1.197864	0.000000
	3.464102	0.000000	0.000000	1.861210
6	1.101021	1.049295	0.000000	0.000000
	1.434878	0.000000	1.197864	0.000000
	3.464102	0.000000	0.000000	1.861210



Summarized, for Rayleigh scattering and $m > 2$ one finds $\Lambda = \kappa$, where κ is the matrix of diffuse extinction coefficients. Also, $Y = Y^T = \kappa^{1/2}$. Now, application of the relationships (cf. section 5.4) $YY^T = A + B$ and $Y^{-T}\Lambda^2Y^{-1} = A - B$ gives $A + B = A - B = \kappa$, or $B = O$ and $A = \kappa$. Since $A = \kappa - F$, it can be concluded that $B = F = O$, so both the forward scattering matrix and the backscattering matrix are equal to the zero matrix. The conclusion is that application of the discrete cosine transform to the azimuthal component of Rayleigh scattering already partly (for $m > 2$) accomplishes the diagonalization of the matrix of extinction and scattering. The reason for this behaviour is that the azimuthal dependence of the Rayleigh phase function (cf. section 9.3) can be expressed by

$$p_R(\psi) = a_0 + a_1 \cos \psi + a_2 \cos 2\psi$$

where ψ is the relative azimuth angle and a_0 , a_1 and a_2 depend on the zenith angles of incidence and observation. Application of the azimuthal DCT gives only a non-zero response for the angular frequencies zero, one and two. For higher frequencies the response is zero, which explains why for $m > 2$ zero matrices F and B are obtained.

6.3 Infinite reflectance matrix

The infinite reflectance matrix describes the diffuse reflectance for infinite optical thickness of the layer and depends only on the eigenvalues and eigenvectors. In this section numerical results are presented for the same cases as for which the eigenvalues and eigenvectors were computed in the previous section. Tables 6.4 to 6.6 give the results for the three cases. They are specified in the form of reflectance factors, i.e. reflectances at the top of the layer relative to those of a white Lambertian surface. The indices i and j refer to the zenith zones of incidence and exitence, respectively and Δk is the absolute difference in the number of azimuth sectors. For the numbering of zones and sectors on the sphere, see Fig. 5.2 in section 5.2. The most right column in the tables gives the partial sums per existent zenith zone and for the three zones added together. For this, the entries for Δk in the range from 1 to 5 have been counted twice, in order to obtain the sum over all twelve azimuth sectors. The totals for each incidence zenith sector give the hemispherical reflectance when these are divided by the total number of segments



Table 6.4 Infinite reflectance matrix for aerosol Haze M at 700 nm and $\omega_A = 1$

<i>i</i>	<i>j</i>	Δk							Sum
		0	1	2	3	4	5	6	
1	1	1.142	1.143	1.141	1.141	1.143	1.147	1.148	13.720
	2	1.045	1.044	1.043	1.049	1.064	1.082	1.089	12.698
	3	0.753	0.753	0.758	0.779	0.817	0.862	0.883	9.574
									35.992
2	1	1.045	1.044	1.043	1.049	1.064	1.082	1.089	12.698
	2	0.989	0.986	0.984	1.006	1.060	1.126	1.157	12.470
	3	0.764	0.757	0.758	0.810	0.939	1.137	1.260	10.826
									35.994
3	1	0.753	0.753	0.758	0.779	0.817	0.862	0.883	9.574
	2	0.763	0.757	0.758	0.810	0.939	1.137	1.260	10.826
	3	0.721	0.700	0.687	0.797	1.185	2.258	3.624	15.599
									35.999

per hemisphere, which is 36. The results are interpreted as follows. The deviations from a Lambertian reflector are moderate for Rayleigh scattering (Table 6.6) and also for aerosol scattering when *i* is 1 or 2. For the aerosol and *i* = 3 (Tables 6.4 and 6.5) the reflectance can become more than three times stronger than for a white Lambertian reflector. This happens for *j* = 3 and $\Delta k = 6$, which corresponds to the case similar to mirror reflection under low incidence angles. However, here the cause is not mirror reflection but the strong forward scattering of aerosols. The same case for Rayleigh scattering (Table 6.6) gives a reflectance of 1.227, which is even less than in the backscattering direction ($\Delta k = 0$) for the same *i* and *j*.

The effect of a small absorption of only one percent of the intercepted flux ($\omega_A = 0.99$) is demonstrated in Table 6.5 by a strong decrease of the infinite reflectances. The strongest absolute decreases are found for small *i* and *j*, i.e. for angles close to the zenith. The hemispherical reflectances for the three incidence zenith zones are 0.561, 0.616 and 0.736 for *i* is 1 to 3, respectively. These figures indicate that a small absorption as expressed in the single scattering albedo may result in a large absorption of the semi-infinite layer, especially at steep incidence angles.



Table 6.5 Infinite reflectance matrix for aerosol Haze M at 700 nm and $\omega_A = 0.99$

<i>i</i>	<i>j</i>	Δk							Sum
		0	1	2	3	4	5	6	
1	1	0.605	0.605	0.603	0.603	0.604	0.607	0.609	7.258
	2	0.581	0.579	0.578	0.583	0.597	0.613	0.620	7.101
	3	0.445	0.445	0.449	0.469	0.505	0.548	0.568	5.845
									20.204
2	1	0.581	0.579	0.578	0.583	0.579	0.613	0.620	7.101
	2	0.585	0.582	0.579	0.599	0.650	0.712	0.742	7.571
	3	0.492	0.486	0.486	0.534	0.658	0.850	0.969	7.489
									22.161
3	1	0.445	0.445	0.449	0.469	0.505	0.548	0.568	5.845
	2	0.492	0.486	0.486	0.534	0.658	0.850	0.969	7.489
	3	0.533	0.512	0.498	0.603	0.979	2.029	3.373	13.148
									26.482

The numerical accuracy of the (N+2)-stream model can partly be checked by considering the hemispherical infinite reflectances in the case of conservative scattering, which should give hemispherical reflectances equal to one. In Table 6.4 the greatest error is found for the aerosol at $i = 1$. In this case the deviation from one is $0.008/36 = 0.00022$, which is acceptable. In Table 6.6 two of the hemispherical sums exceed 36, thus giving hemispherical reflectances greater than one, which is physically impossible. However, the sums have been calculated afterwards from the printed computer output which listed only the individual reflectances in three decimal places, and when the effect of rounding errors on the hemispherical sums is taken into account, it is concluded that the deviations from one are all too small to be explained by numerical inaccuracy of the algorithm which calculates the infinite reflectance matrix. Later in this chapter (section 6.4) more definite statements on the issue of numerical accuracy are made on the basis of model simulations in which the layer absorption has been explicitly calculated from hemispherical reflectance and transmittance for specular incidence within the computer program, so that rounding errors due to printing can play no role.



Table 6.6 Infinite reflectance matrix for Rayleigh scattering

<i>i</i>	<i>j</i>	Δk							Sum
		0	1	2	3	4	5	6	
1	1	1.107	1.103	1.091	1.078	1.065	1.057	1.054	12.949
	2	1.075	1.067	1.047	1.024	1.005	0.995	0.992	12.343
	3	0.933	0.924	0.901	0.881	0.872	0.873	0.875	10.710
									36.002
2	1	1.075	1.067	1.047	1.024	1.005	0.995	0.992	12.343
	2	1.098	1.080	1.039	0.997	0.973	0.965	0.965	12.171
	3	1.048	1.021	0.961	0.916	0.911	0.935	0.949	11.485
									35.999
3	1	0.933	0.924	0.901	0.881	0.872	0.873	0.875	10.710
	2	1.048	1.021	0.961	0.916	0.911	0.935	0.949	11.485
	3	1.318	1.251	1.112	1.029	1.067	1.173	1.227	13.809
									36.004

6.4 Hemispherical reflectance and transmittance for specular incidence

In the (N+2)-stream model the diffuse reflectance and transmittance for specular incidence are described by means of the vectors ρ_{sd} and τ_{sd} . For N equal to 72 these vectors consist of 36 elements, which describe the distribution of the reflected and transmitted radiation over the 36 angular segments in the upper and lower hemispheres, respectively. Summation of all elements of each vector gives the hemispherical reflectance and transmittance for specular incidence. Since the sum of both, plus the direct transmittance, gives the fraction of the incident flux exiting from the layer, one minus this fraction is the fraction absorbed, which is called absorbance. This quantity can be used to check the numerical accuracy of the model if calculations are performed for the case of conservative scattering and very high optical thickness. In that case the numerically determined absorbance should still be zero.

In section 6.2 it was shown that the smallest eigenvalue obtained for aerosol



scattering and $\omega_A = 1$ is equal to 0.000018, whereas it should be zero. The consequence of this numerical inaccuracy for the absorbance of an optically very thick layer ($\log b = 5.6$) has been investigated for solar zenith angles of 15, 45 and 75 degrees. The result is that the calculated absorbance becomes 0.000127, 0.000104 and 0.000064, respectively. For Rayleigh scattering even at $\log b = 6.1$ no absorption was found when represented to six decimal places. From this it is concluded that the inaccurate calculation of the smallest eigenvalue indeed leads to a numerical loss of energy, but even at unrealistically high optical thickness the numerically induced layer absorption associated with this can be considered negligible.

For aerosol scattering, the relationships between single scattering albedo, smallest eigenvalue ($\lambda_{\min}$), saturation optical thickness b_{sat} , and layer absorbance for a solar zenith angle of 45 degrees have been investigated by means of the model and the results are presented in Table 6.7. Here the saturation optical thickness b_{sat} is defined as the optical thickness above which the layer absorbance (to six decimal places) does not increase any more. This is also the point where the hemispherical transmittance has reduced to less than 10^{-6} . The layer absorbance at the saturation point is symbolized as α_{sat} .

Table 6.7 Relationships between single scattering albedo, minimum eigenvalue, saturation optical thickness and saturation absorbance

ω_A	$\lambda_{\min}$	$\log b_{\text{sat}}$	α_{sat}
0.99999	0.002741	3.7	0.015516
0.9999	0.008669	3.2	0.048222
0.999	0.027424	2.7	0.144482
0.99	0.087075	2.3	0.388013
0.9	0.285863	1.8	0.783850
0.5	0.720344	1.3	0.969731
0.1	1.034135	1.1	0.996854

From these results one may observe the following:

- For weak absorption $\lambda_{\min} \approx 0.87 \sqrt{(1-\omega_A)}$ and $\alpha_{\text{sat}} \approx 4.9 \sqrt{(1-\omega_A)}$. Similar relationships can be derived analytically from Kubelka-Munk (two-stream) theory. They explain the great sensitivities to the single scattering albedo when this is close to one. Theoretically, at $\omega_A = 1$ this sensitivity is infinite because of the square root function.
- The product $\lambda_{\min} b_{\text{sat}}$ is approximately constant and amounts to about 15. However, for Rayleigh scattering $\lambda_{\min}$ was found to be zero, but still saturation occurred at $\log b = 6.1$, so this relationship does no longer hold when $\lambda_{\min}$ becomes very small.

The above results indicate that when very small amounts of absorbing material are added to an optically thick layer consisting of purely scattering material, the absorption of incident radiation by the layer as a whole can already become considerable. For instance, when only 0.001 % (in units of optical thickness) of absorbing material is added, the absorption by the layer of solar radiation under 45 degrees incidence becomes already more than 1.5 % . Also, 50 % of absorbing material gives 97 % layer absorbance, or 3 % reflectance, which already indicates a virtually black surface.

For a large range of optical thicknesses, spanning eleven orders of magnitude (from $b = 10^{-5}$ to $b = 10^6$), calculations of the direct transmittance, the diffuse hemispherical reflectance, the diffuse hemispherical transmittance and the layer absorbance have been carried out for aerosol scattering as well as Rayleigh scattering and for different solar zenith angles. For the range $b = 10^{-3}$ to 10^4 some results are presented in Figures 6.1 to 6.5. In all Figures a logarithmic scale for b is applied because of the large range, and linear scales for reflectances and transmittances. In order to show the division of the radiation over transmittance, reflectance and absorbance, the left vertical scale is used for the total and hemispherical diffuse transmittance, whereas the right vertical scale runs in opposite direction and is used for the hemispherical diffuse reflectance. Fig. 6.1 shows the results for the water Haze M aerosol at 700 nm, a single scattering albedo of one and a solar zenith angle of 15 degrees. In this case there is no absorption of enough significance to be plotted, so the sum of hemispherical reflectance and total transmittance is one everywhere. The course of the optical quantities as a function of the optical thickness b can be characterized as follows:



- For small b , say less than 0.1, the direct transmittance decreases linearly, and the diffuse reflectance and transmittance increase linearly. This follows from theoretical considerations, cf. section 5.5. In this region, the hemispherical diffuse transmittance is about twenty times stronger than the reflectance, which is explained by the strong forward scattering of the aerosol.
- For higher b the direct transmittance decreases exponentially (actually it does for all b), and becomes negligible at $b = 6$. The intercepted fraction is divided over diffuse reflectance and transmittance, so the sum of these increases up to $b = 6$, where it becomes one. The diffuse reflectance is a monotonously increasing function of b , which is first linear, then approximately logarithmic and finally (for $b > 10$) slower than logarithmic. As a result of this, the diffuse transmittance has a maximum, which in this case is located at $b = 2.5$ and amounts to 73 %.
- The curve for the diffuse transmittance as a function of $\log b$ is highly symmetrical about the vertical line through the maximum. This suggests that for conservative scattering at very high optical thickness the diffuse transmittance is proportional to the inverse, b^{-1} . This kind of behaviour can be confirmed theoretically for the four-stream model (i.e. $N = 2$), so it seems plausible that this is a general feature.

Figures 6.2 and 6.3 are similar to Fig. 6.1 except for the solar zenith angle, which is 45 and 75 degrees, respectively. From these it is concluded that for larger solar zenith angles the direct transmittance decreases more rapidly, so that the diffuse reflectance and transmittance increase faster as a function of b for small b . The maximum diffuse transmittance is also reached at smaller b , and is lower (60 % for $\theta_s = 75$ degrees). The curve for the diffuse transmittance becomes less symmetric with respect to the maximum.

Figure 6.4 shows the results obtained for Rayleigh scattering with $\theta_s = 45$ degrees. Compared with Fig. 6.2 it is seen that the general features are the same, but the diffuse reflectance is much higher for Rayleigh scattering, while the diffuse transmittance is much lower. The maximum diffuse transmittance in Fig. 6.4 is only 35.5 % instead of 68.5 % in Fig. 6.2. These differences can be explained by the Rayleigh phase function which is equal for forward and backward scattering. In Fig. 6.4 indeed one sees about equal diffuse reflectances and transmittances for small b .

Fig. 6.5 shows the effect of a small fraction of absorption ($\omega_A = 0.99$) under conditions otherwise equal to those of Fig. 6.2. Here it appears that for b

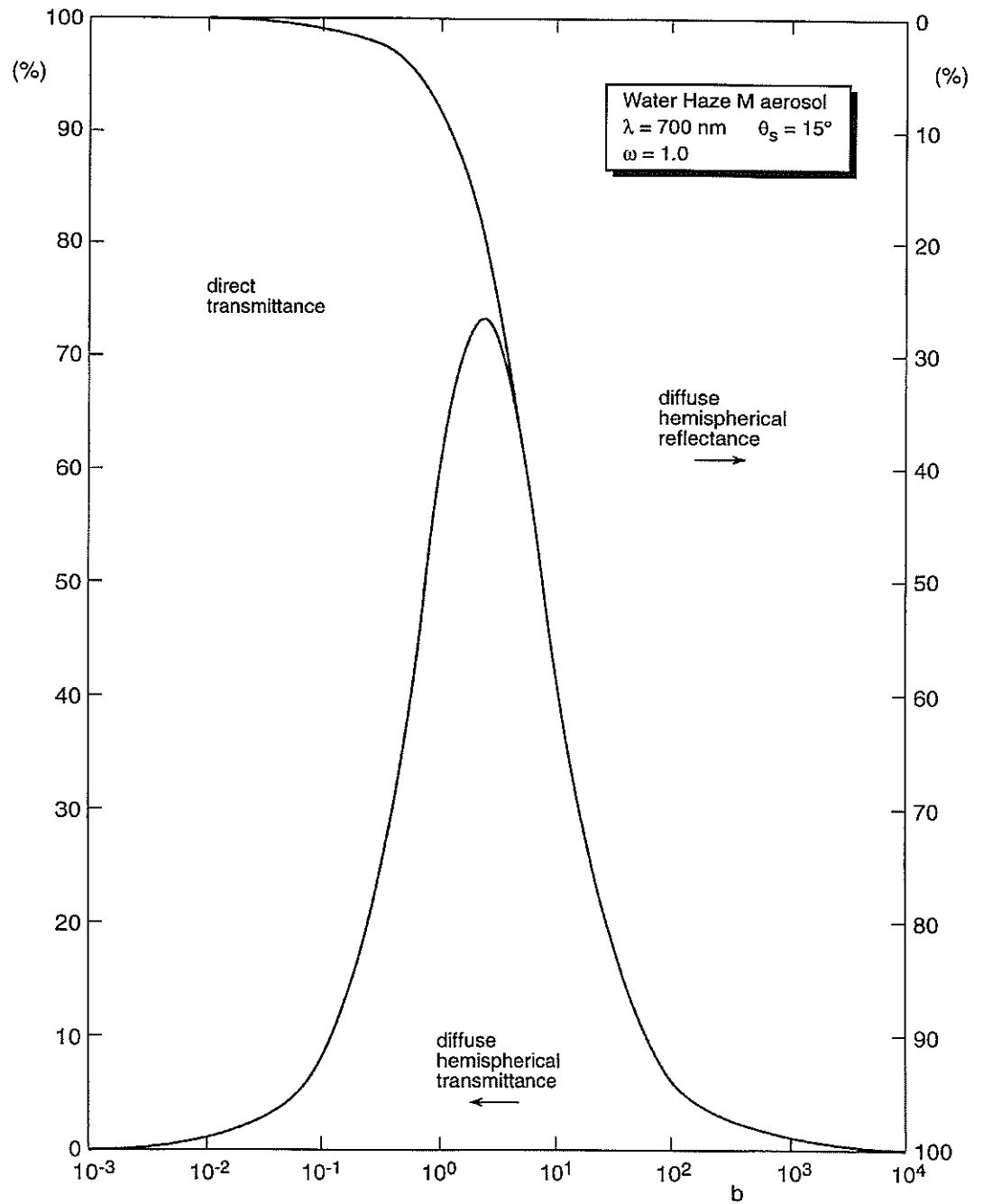


Fig. 6.1 Radiation budget as a function of optical thickness for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 15^\circ$

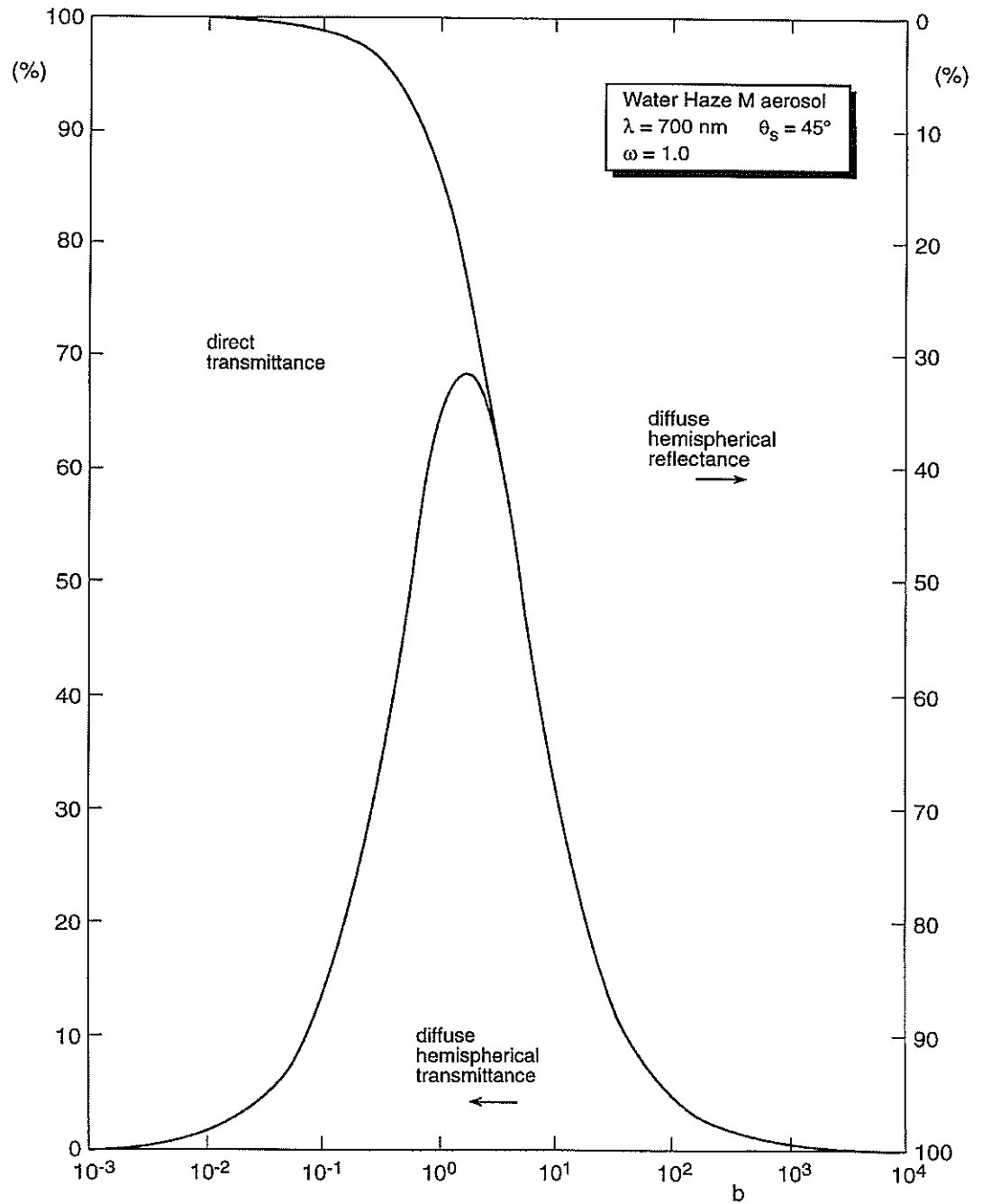


Fig. 6.2 Radiation budget as a function of optical thickness for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$

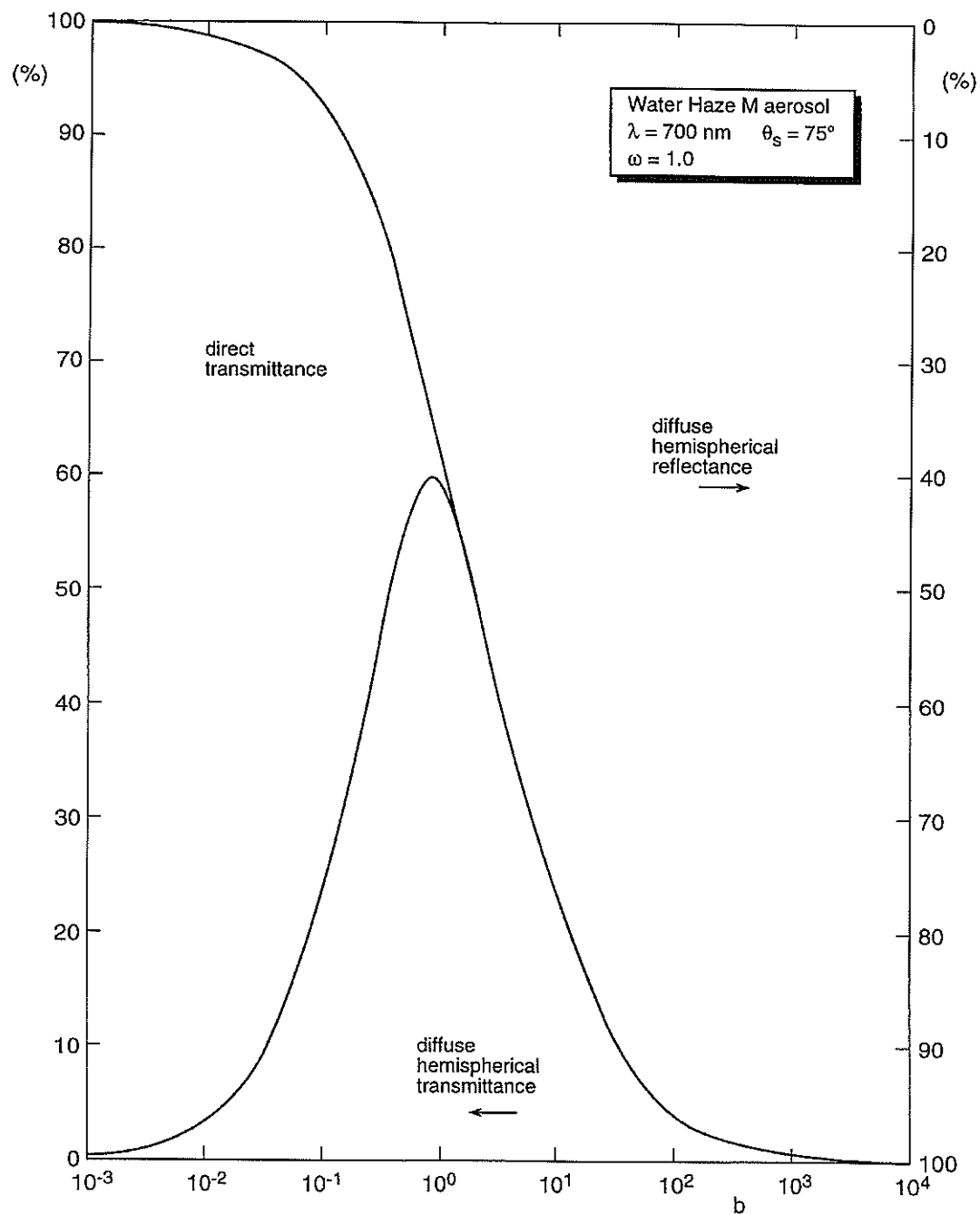


Fig. 6.3 Radiation budget as a function of optical thickness for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 75^\circ$

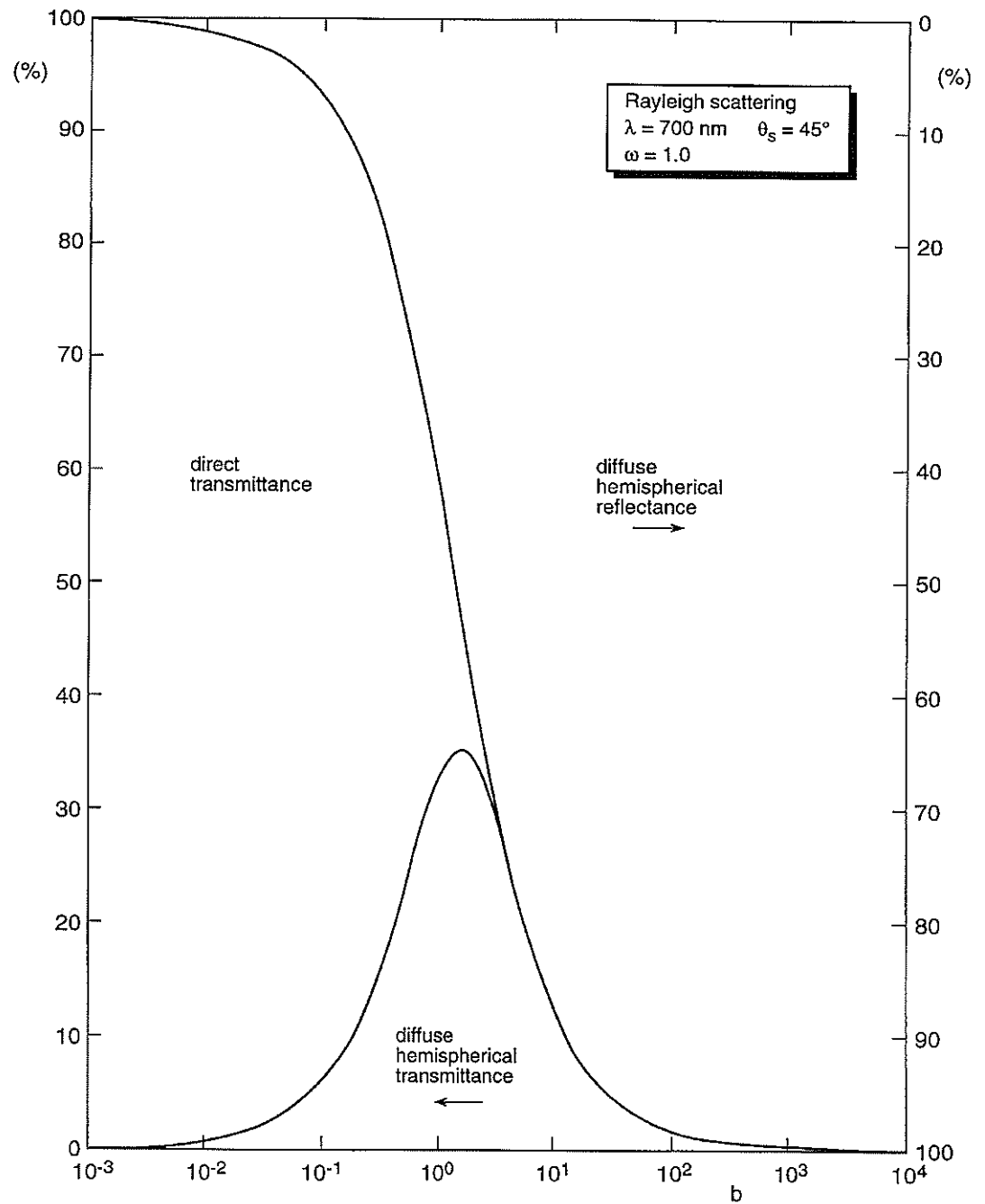


Fig. 6.4 Radiation budget as a function of optical thickness for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$

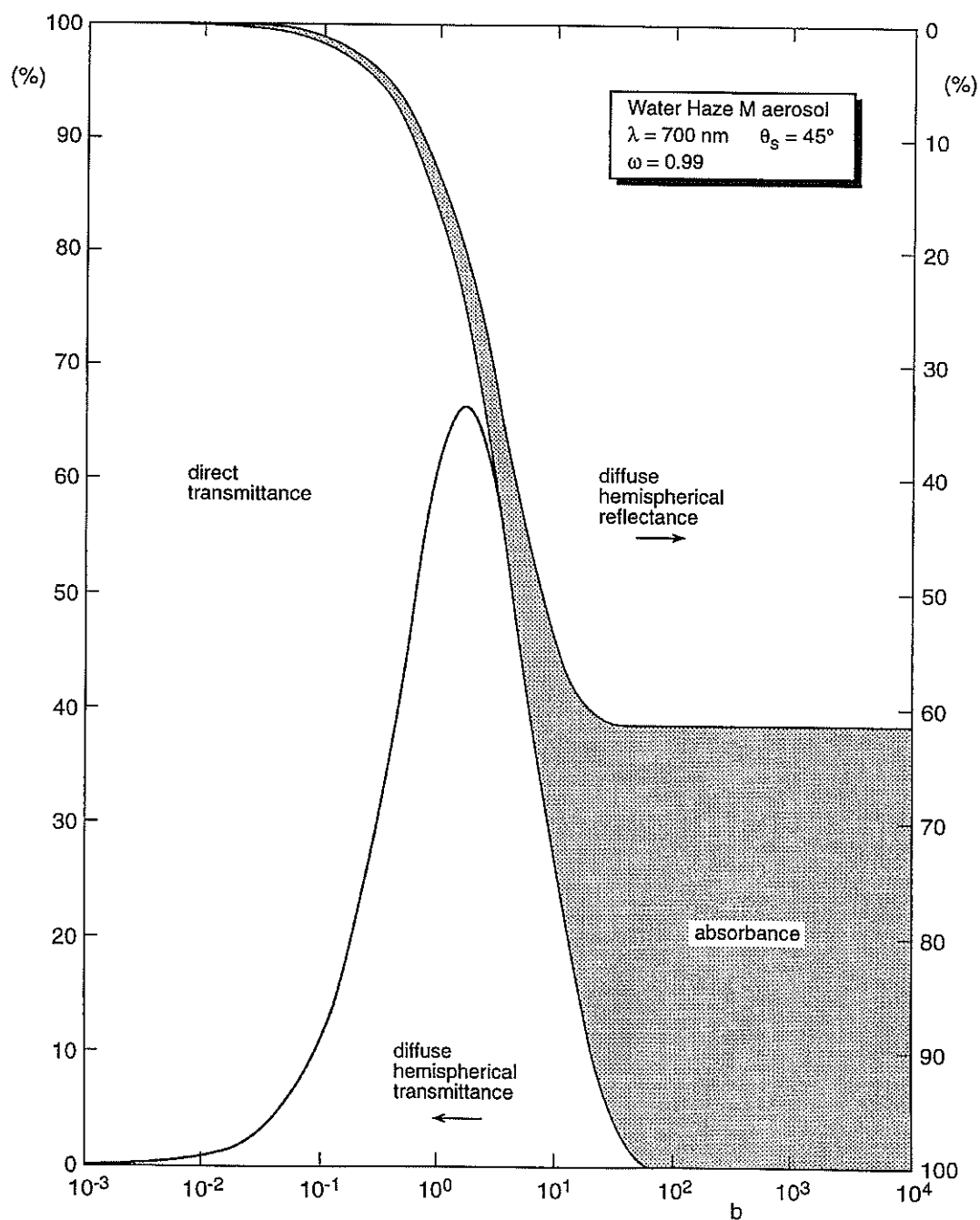


Fig. 6.5 Radiation budget as a function of optical thickness for water Haze M aerosol at 700 nm for $\omega = 0.99$ and $\theta_s = 45^\circ$



less than 1 there is not much difference, except for a small layer absorbance, increasing to about 2 % at $b = 1$. However, at larger b the layer absorbance increases rapidly to almost 39 % and then saturates at $b = 100$. One may also observe that the diffuse reflectance saturates already at $b = 30$. Here the diffuse transmittance is still about 4 % and reduces to zero at $b = 100$. Above this optical thickness everything has become constant.

The results presented so far demonstrate that the $(N+2)$ -stream model can be applied to radiation budget calculations for a very wide range of optical thicknesses, spanning eleven orders of magnitude, and for various types of scattering phase functions, including highly anisotropic ones, such as those for aerosol scattering.

6.5 Bidirectional reflectance and transmittance

In this section the performance of the $(N+2)$ -stream model (for $N = 72$) is investigated by means of simulation of the bidirectional reflectance and transmittance for a wide range of optical thicknesses and for the phase functions of aerosol and Rayleigh scattering. The calculations have been limited to the principal plane (i.e. relative viewing azimuth either 0 or 180 degrees) as this gives the greatest directional variations, due to the fact that in this case the complete range of possible scattering angles is encountered. The viewing zenith angle has been varied in steps of 2.5 degrees, which is small enough to resolve the angular variation of the aerosol scattering phase function in the so-called glory region (close to exact backward scattering) and the aureole region (close to exact forward scattering). The optical thickness has been varied exponentially from $b = 2^{-6}$ to 2^{12} , with a step factor of four. The results have been plotted on logarithmic scales for the bidirectional reflectance and transmittance. The quantities shown are $r_{so} = (N/2) \rho_{so}$ and $t_{so} = (N/2) \tau_{so}$, where ρ_{so} and τ_{so} are the bidirectional reflectance and transmittance in terms of flux-equivalent radiance as defined in chapter 5. Superscripts s and d are used to indicate contributions due to single scattering and multiple scattering, respectively.

Fig. 6.6 shows the single scattering contribution to the bidirectional reflectance for water Haze M aerosol at 700 nm, a single scattering albedo of one, and a solar zenith angle of 45 degrees. This illustration is included only to show the difference with the multiple scattering contribution, which is

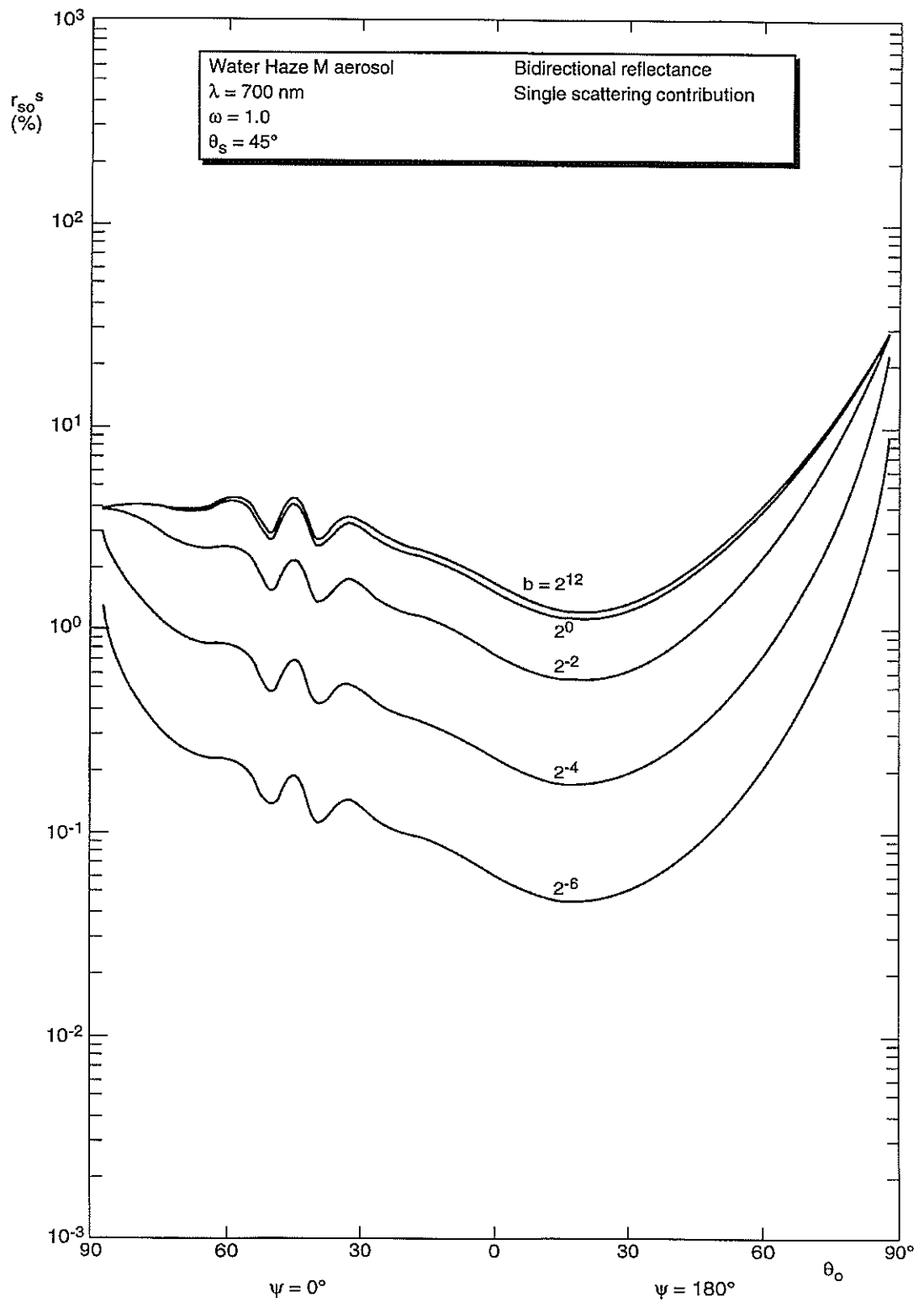


Fig. 6.6 Bidirectional reflectance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Single scattering contribution



presented in Fig. 6.7. It cannot be used to demonstrate the performance of the $(N+2)$ -stream model, since the single scattering contribution can easily be calculated in any model by means of a stand-alone and exact formula. Nevertheless, Fig. 6.6 serves also to illustrate the most prominent features of single scattering for aerosols as a function of the viewing angle and the optical thickness. A striking feature in Fig. 6.6 is the ripple in the neighbourhood of the anti-solar direction ($\theta_o = 45^\circ; \psi = 0^\circ$). This oscillatory behaviour is a property of the phase function of non-absorbing spherical water droplets, especially when these have sufficient size in comparison to the wavelength of the radiation. The Haze M particle size distribution (Deirmendjian, 1969) has relatively many of these particles.

The other feature of aerosol scattering is the strong forward scattering relative to backscattering. In Fig. 6.6 this is manifested by the strong increase of the bidirectional reflectance with increasing viewing angle for $\psi = 180^\circ$. Close to $\theta_o = 90^\circ$ this effect gives already bidirectional reflectances of more than 10 % for optically thin layers.

As a function of the optical thickness b the single scattering contribution to the bidirectional reflectance increases almost linearly for small b , given the approximately fourfold increase of it when b is increased from 2^{-6} to 2^{-4} . However, saturation of this contribution occurs already for moderate values of b , especially for large viewing zenith angles. But even at zero viewing zenith angle saturation takes place just above $b = 1$, given the fact the curves for $b = 1$ and $b = 2^{12}$ (4096) differ only about 10 % at this viewing angle.

For the same conditions as in Fig. 6.6, Fig. 6.7 gives the multiple scattering contribution, except that a few more optical thicknesses have been included and that for $b = 0.25$ also the results for $N = 2$ (four-stream model) are shown. The most striking feature of Fig. 6.7 is the much larger dynamic range of this contribution when compared to the single scattering contribution for the same range of optical thicknesses. For the latter this was about a factor of twenty, for the former it is a factor of almost 100,000. At small b the directional pattern is roughly similar to that of the single scattering contribution, except that the ripples in the glory region have been washed out. As predicted in chapter 5, for small b this contribution increases quadratically as a function of the optical thickness, and this is confirmed in Fig. 6.7, given the approximately 16-fold increase when b increases from 2^{-6} to 2^{-4} . Due to this effect, already at $b = 0.25$ both contributions are of the same order of magnitude. For this optical thickness, which is also quite

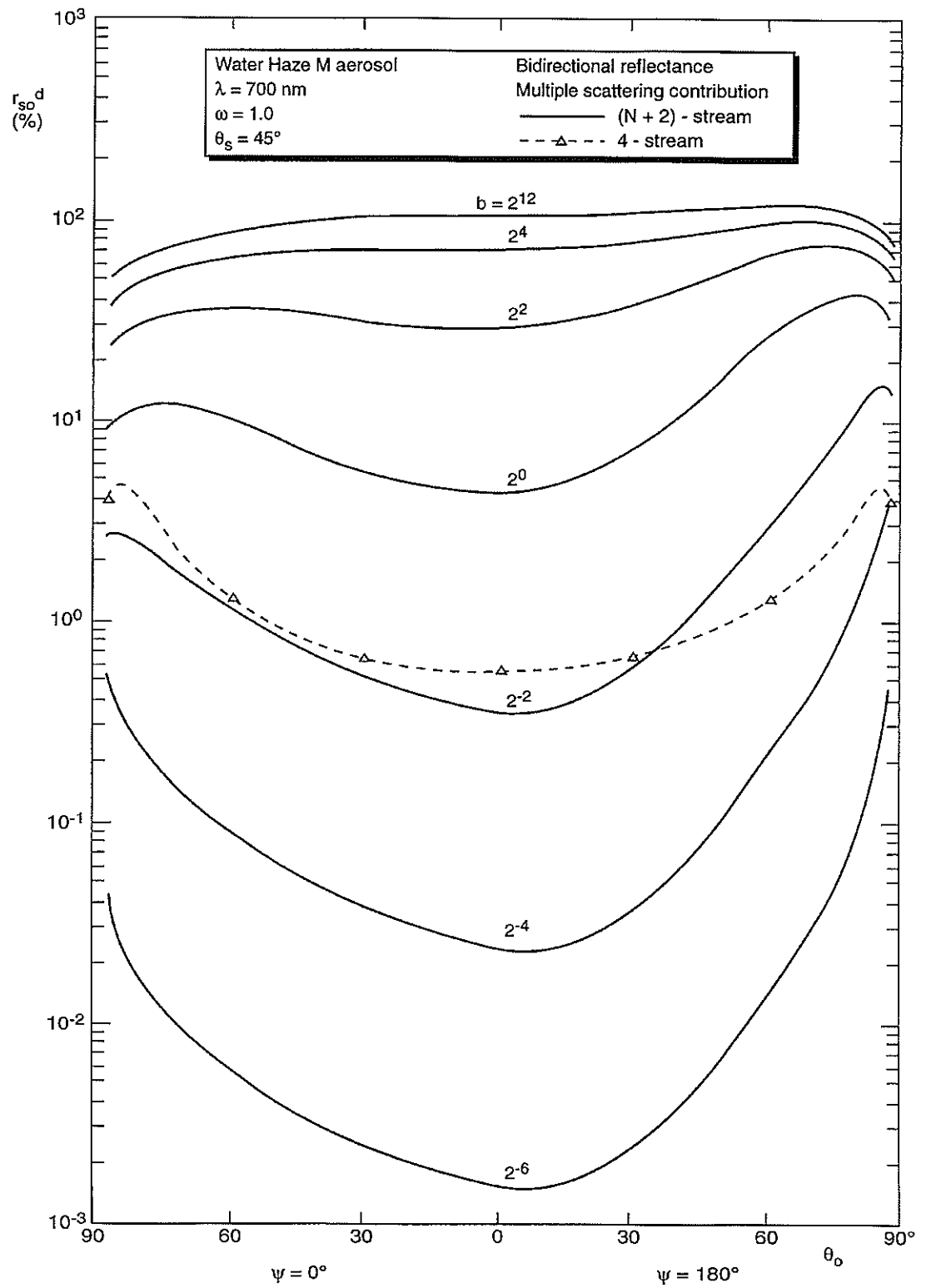


Fig. 6.7 Bidirectional reflectance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Multiple scattering contribution



representative for atmospheric conditions under which optical remote sensing can successfully be executed, the results obtained for the four-stream model are also given. It appears that the correct order of magnitude is obtained, but the directional pattern differs considerably from the $(N+2)$ -stream result. Especially in the forward scattering region ($\psi = 180^\circ$) the differences between the two results are substantial, with maximum relative deviations of about a factor of three at large zenith viewing angles. It should be noted here that the directional pattern resulting from the four-stream model is azimuth-independent, and since for the $(N+2)$ -stream model the greatest directional variations are located in the principal plane, this is probably a worst case.

From Fig. 6.7 it appears also that the multiple scattering contribution still grows considerably above $b = 1$, and saturation is reached at a much higher optical thickness than was found for the single scattering contribution. At saturation (for $b = 4096$) values of about 100 % are found, with a decrease for large viewing zenith angles ("limb darkening").

In Fig. 6.8 the sum of the single and multiple scattering contributions is presented, thus giving the total bidirectional reflectance. From this it is seen that for small b the result is mainly determined by the single scattering contribution, whereas above $b = 0.25$ the multiple scattering contribution dominates. The ripple in the glory region, which is a single scattering effect, seems to become gradually smaller with increasing optical thickness, but this is due to the logarithmic scale, which expresses relative differences. In absolute sense the ripple becomes even stronger, since in Fig. 6.6 it is virtually constant on a logarithmic scale while its level increases, which means that it is increasing with b in absolute sense. Again for $b = 0.25$ also the result for the four-stream model is shown. The relative difference with the $(N+2)$ -stream result is about 20 % in the region around zero viewing zenith angle, and becomes about 50 % for large viewing angles in the forward scattering part of the principal plane. From this it may be concluded that for accurate calculations of the atmospheric influence on optical remote sensing imagery a four-stream model is inadequate. However, this does not make the four-stream model completely useless, since there will not be many cases in which the accuracy of the model input parameters is so high that use of a more accurate model would be justified.

For the bidirectional transmittance a similar analysis is presented in Figures 6.9 to 6.12. Fig. 6.9 shows the single scattering contribution to the bidirectional transmittance for the same conditions as in Fig. 6.6. The most striking feature here is the high forward scattering peak at $\theta_o = 135^\circ$ and $\psi = 180^\circ$,

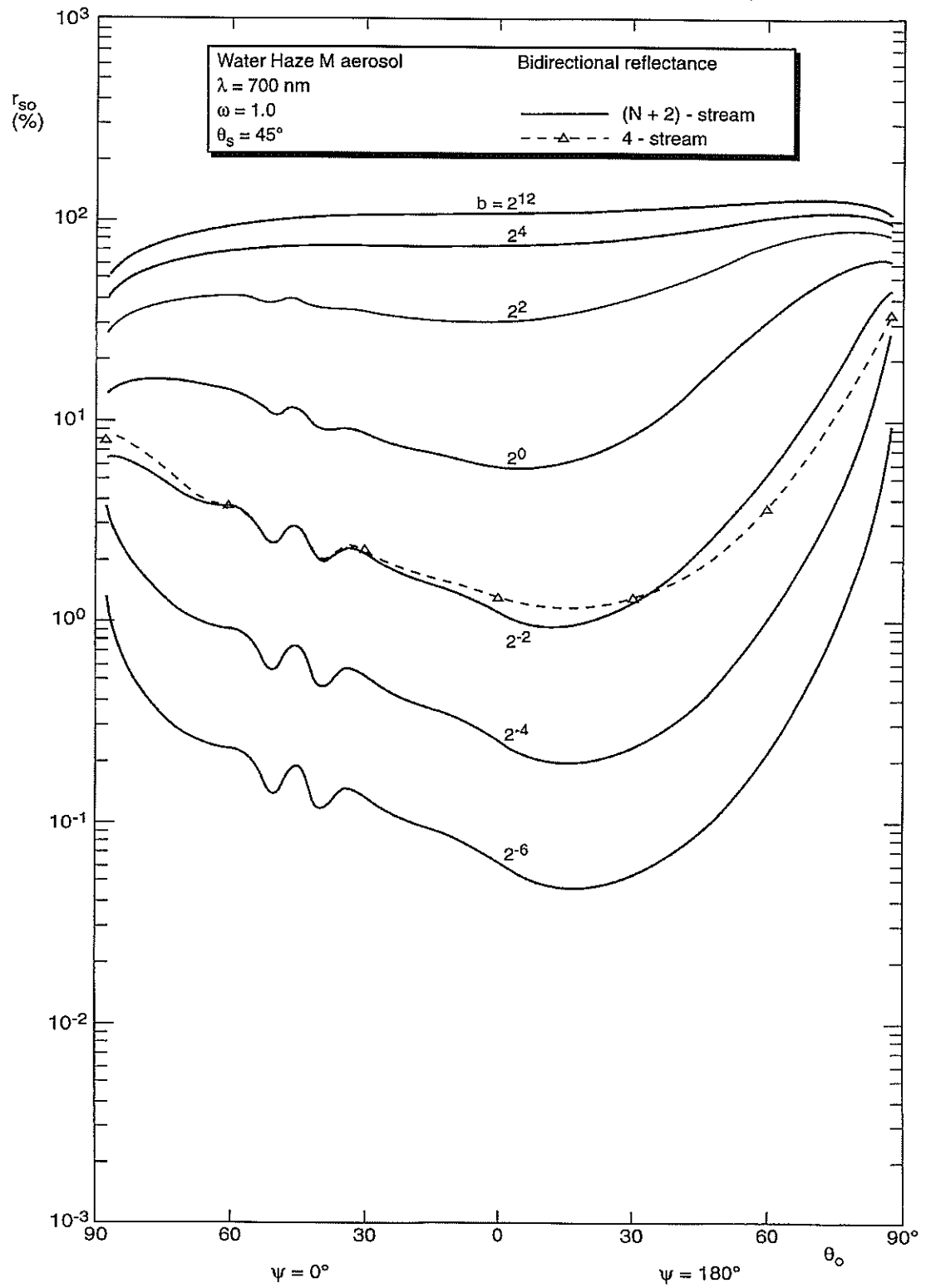


Fig. 6.8 Total bidirectional reflectance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$



causing a very large dynamic range spanning a factor of about 500. For small b and not too large θ_o the directional pattern increases linearly with b , like was found in Fig. 6.6. However, for most viewing angles the single scattering contribution to the bidirectional transmittance becomes maximum at $b = 1$ and then decreases rapidly. At $b = 4$ it has reached a level of the same order as for $b = 2^{-6}$ and at $b = 16$ this contribution was already too small to be drawn in this Figure. Fig. 6.10 shows the multiple scattering contribution. Like was found for the bidirectional reflectance, also in this case the directional pattern for small b has many of the features of the associate single scattering contribution. The forward scattering peak in the aureole region has been washed out considerably, but its influence is still visible. The increase as a function of b for small b is again quadratic. It slows down first at viewing angles close to the horizon ($\theta_o = 90^\circ$) and later also at more steep viewing angles and a maximum is reached at $b = 1$ to 4, depending on the viewing direction. At $b = 4$ still a faint forward scattering peak is visible, but at $b = 16$ it has disappeared. Contrary to the single scattering contribution, the multiple scattering contribution at $b = 16$ is still relatively high, namely about 20 %, with a maximum of 35 % at the zenith ($\theta_o = 180^\circ$). For $b = 0.25$ also the four-stream result is shown and in this case the differences with the (N+2)-stream result are dramatic, with deviations of a factor of about 4 at many viewing angles. Only in the zenith direction the deviation is relatively small, because both curves intersect close to this point.

The sum of both contributions to the bidirectional transmittance is shown in Figures 6.11 and 6.12. Figure 6.11 shows the results for small to medium optical thickness, which are still dominated by the single scattering contribution, especially in the aureole region. The increasing bidirectional transmittance towards viewing angles close to 90° and below $b = 1$ is responsible for the brightness of the sky just above the horizon which often can be observed also in reality under clear conditions. For more hazy atmospheres ($b = 1$) this brightening of the region above the horizon is no longer present because other parts of the sky have become about equally bright. In this case the forward scattering peak is still strong and has broadened somewhat. For $b = 0.25$ again the four-stream result is also shown. The relative deviation with the (N+2)-stream result is very small in the aureole region, due to the fact that single scattering dominates here. However, in the region where the bidirectional transmittance is small the relative difference becomes more than a factor of three.

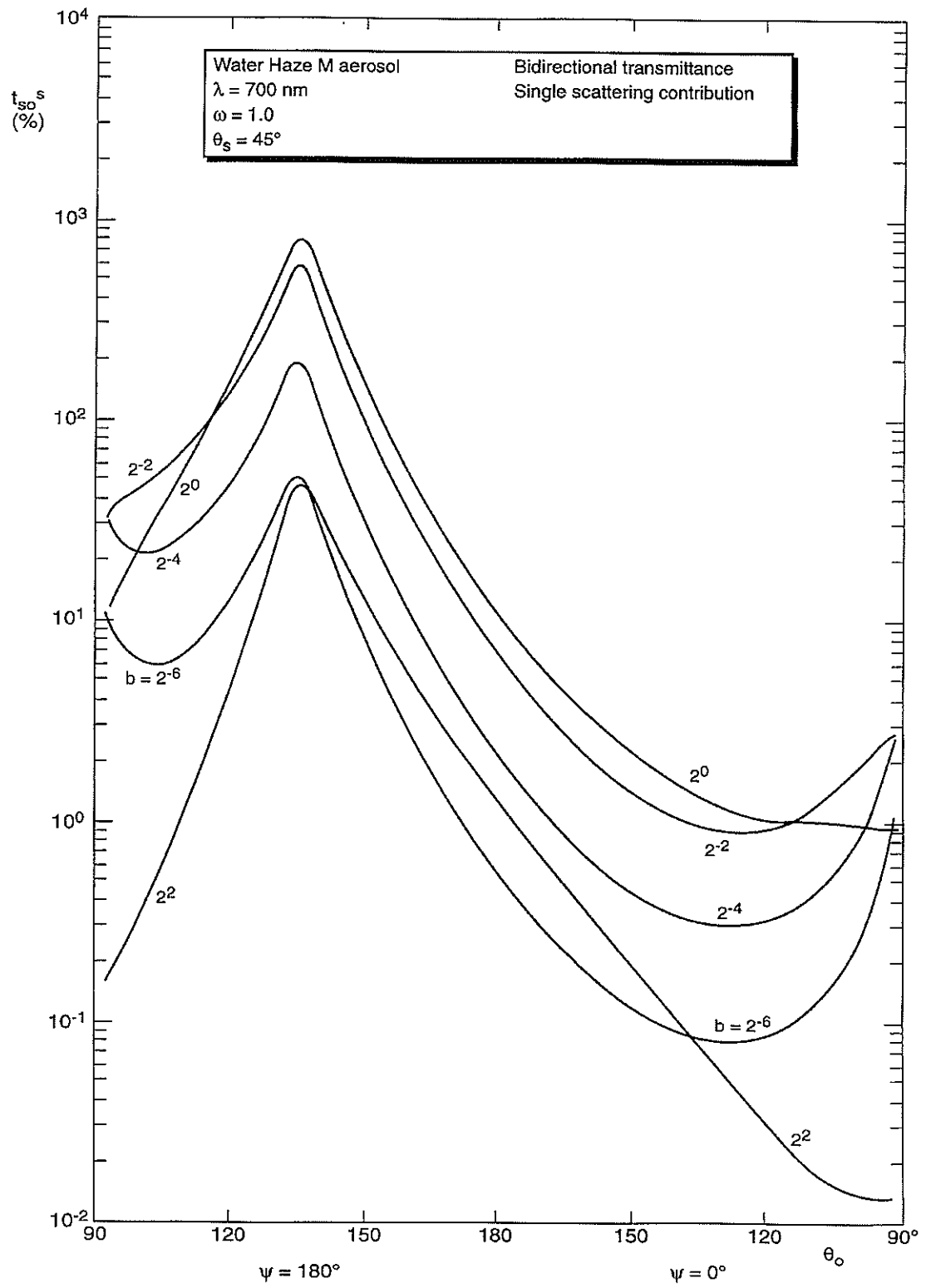


Fig. 6.9 Bidirectional transmittance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Single scattering contribution

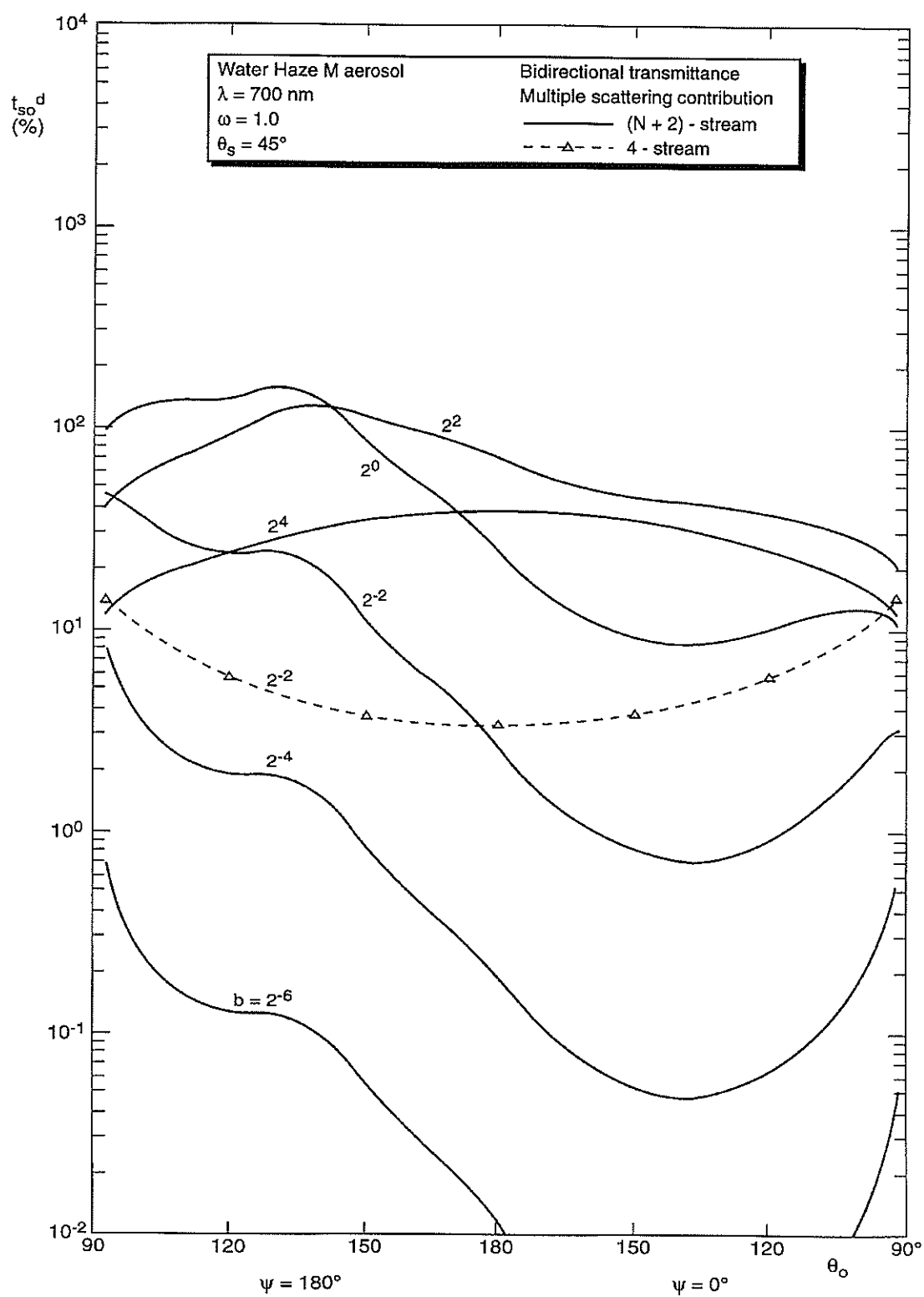


Fig. 6.10 Bidirectional transmittance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Multiple scattering contribution

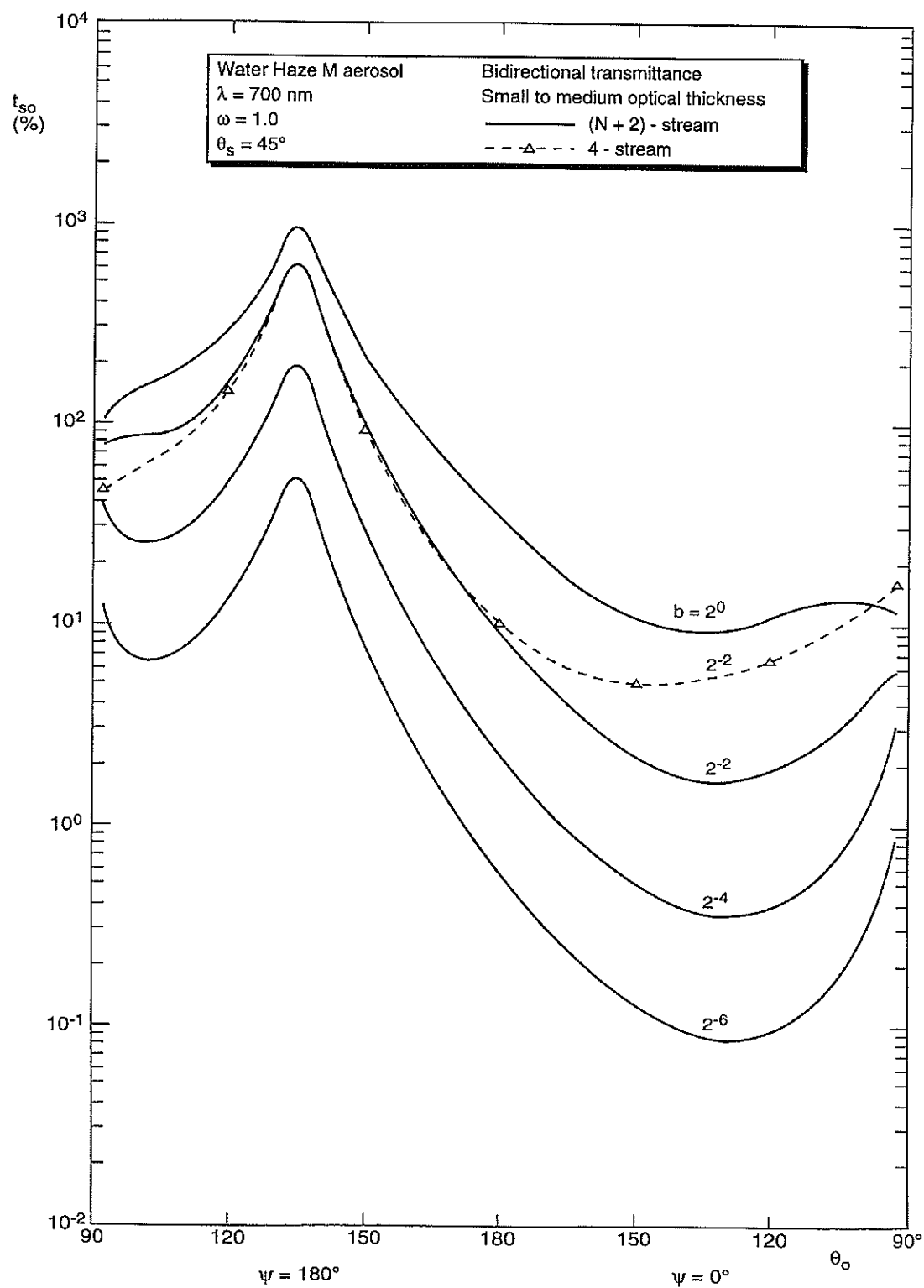


Fig. 6.11 Total bidirectional transmittance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Small to medium optical thickness

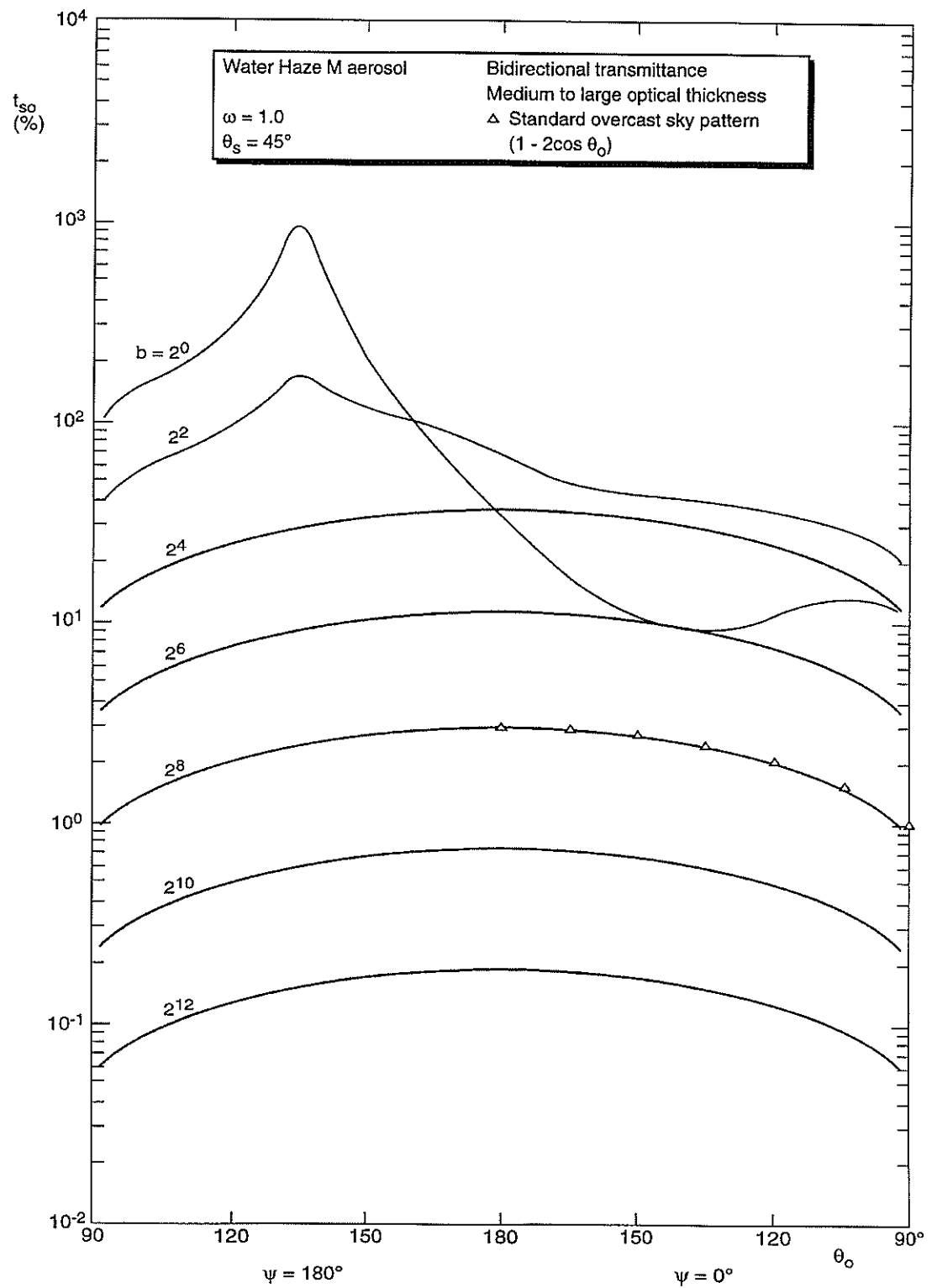


Fig. 6.12 Total bidirectional transmittance in the principal plane for water Haze M aerosol at 700 nm for $\omega = 1$ and $\theta_s = 45^\circ$. Medium to large optical thickness



Fig. 6.12 shows the bidirectional transmittance for medium to large optical thickness. At $b = 4$ a broadened forward scattering peak is still present, but at $b = 16$ it has disappeared and a dome-shaped directional pattern remains, with a flat maximum at the zenith direction and a decreasing bidirectional transmittance towards the horizon. This pattern does not change any more with increasing optical thickness. Only the level decreases inversely proportional to b , as manifested by the parallel displacement of the whole pattern by a factor of about 0.25 for each fourfold increase of the optical thickness. This is consistent with the behaviour of the hemispherical diffuse transmittance for high optical thickness, as noted in the previous section. The directional pattern found for high optical thickness is azimuth-independent and closely resembles the one for the brightness of a so-called Standard Overcast Sky, which (for our definition of the viewing zenith angle) is given by the simple expression $1 - 2 \cos \theta_v$. Den Dulk (1989) applied this sky radiance distribution as input for his canopy reflectance models. It appears to originate from an empirical relation found by Moon & Spencer (1942) and has been verified by Grace (1971). For comparison, in Fig. 6.12 this brightness pattern is indicated by the triangles. There is a remarkable resemblance between the model simulation and the empirical relationship for the radiance pattern of an overcast sky, in spite of the fact that cloud particles have a phase function which is different from that of aerosols, which suggests that perhaps this pattern is independent of the phase function and will always result for very high optical thickness and conservative scattering.

Figures 6.13 to 6.15 present the single, multiple and total contributions to the bidirectional reflectance for the Rayleigh scattering phase function, respectively, for the same range of optical thicknesses and a solar zenith angle of 45° . The directional patterns in this case are much smoother than for the aerosol, but otherwise similar features are found. In general, higher bidirectional reflectances are found for Rayleigh scattering, due to the relatively stronger backscattering in the associate phase function.

Bidirectional transmittances for Rayleigh scattering are shown in Figures 6.16 (small to medium optical thickness) and 6.17 (medium to high optical thickness). Again similar features are found as for aerosol scattering in Figures 6.11 and 6.12, but at high optical thicknesses Rayleigh scattering (compare Figs. 6.17 and 6.12) produces about a fourfold weaker bidirectional transmittance at the same optical thickness, which can be attributed to the strong forward scattering of aerosols. Like in Fig. 6.12, the sky radiance

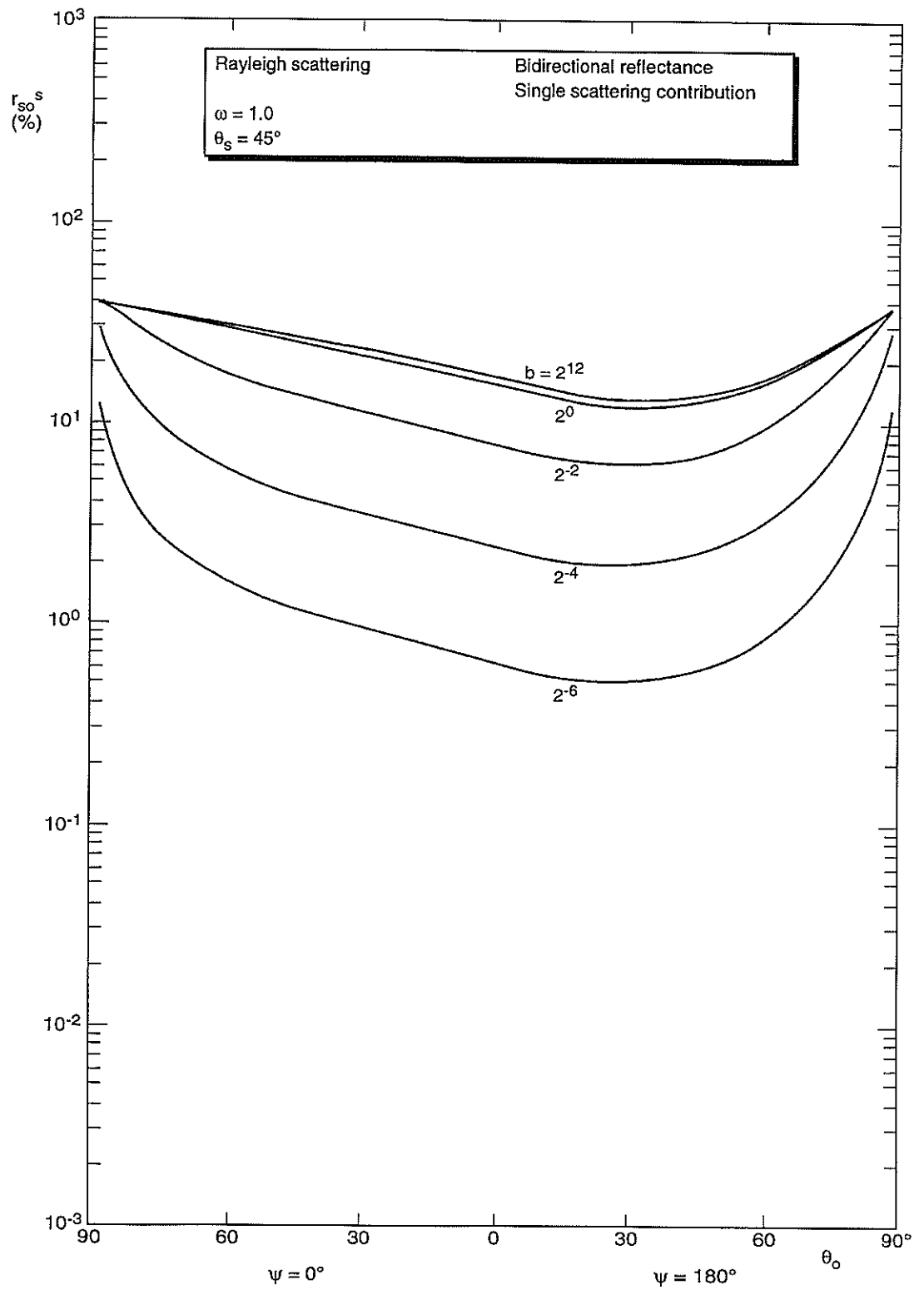


Fig. 6.13 Bidirectional reflectance in the principal plane for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$. Single scattering contribution

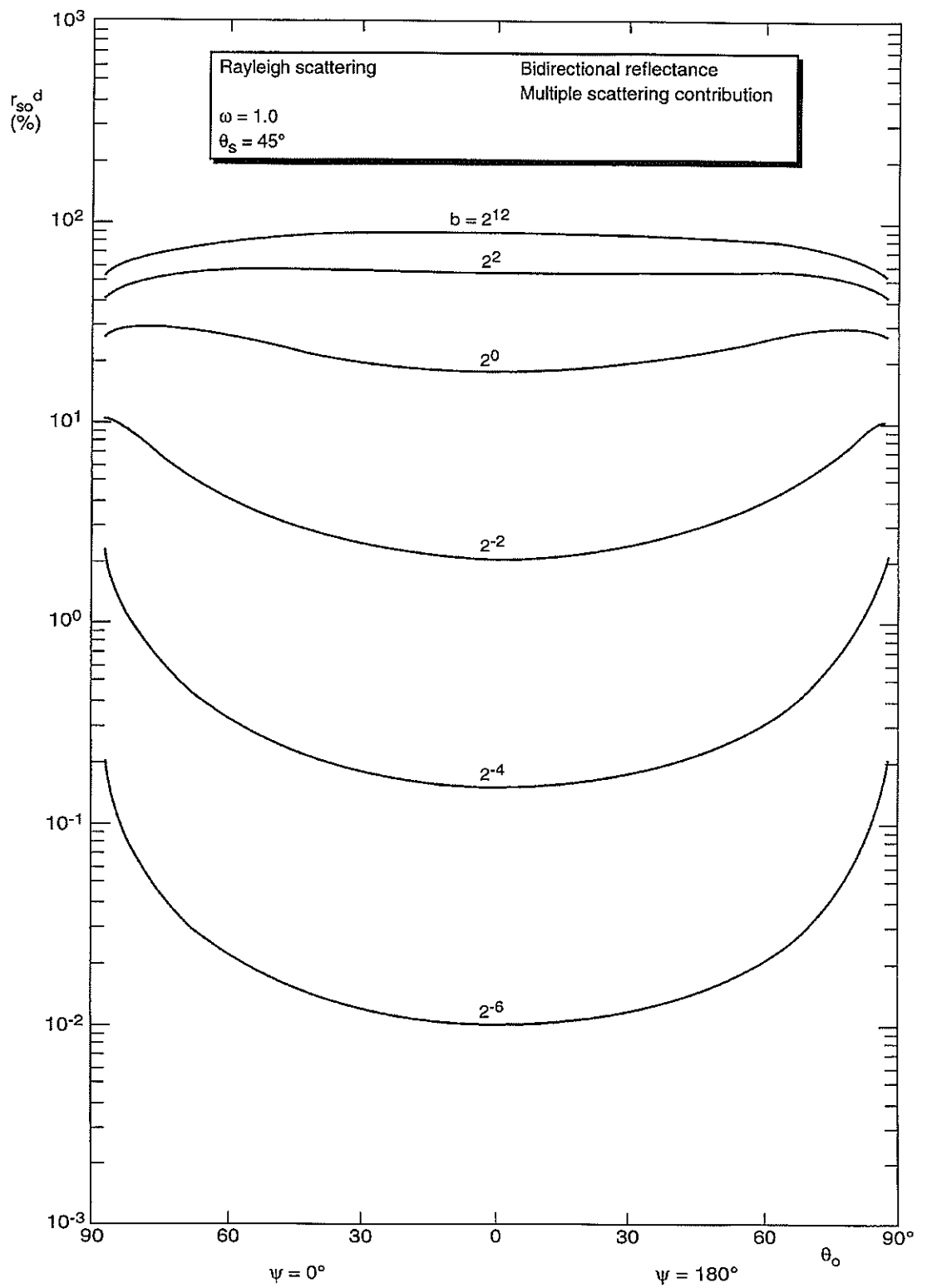


Fig. 6.14 Bidirectional reflectance in the principal plane for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$. Multiple scattering contribution

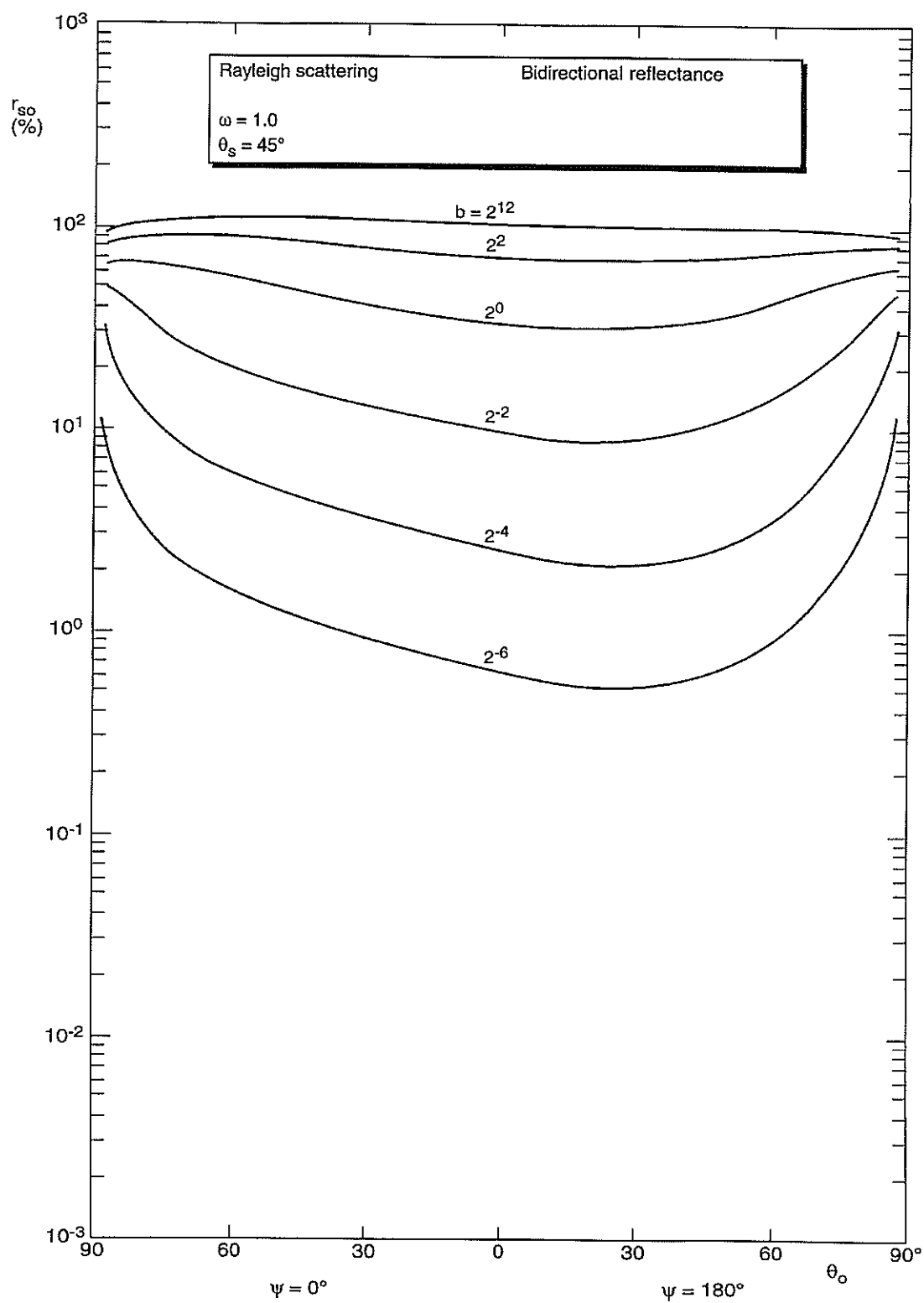


Fig. 6.15 Total bidirectional reflectance in the principal plane for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$

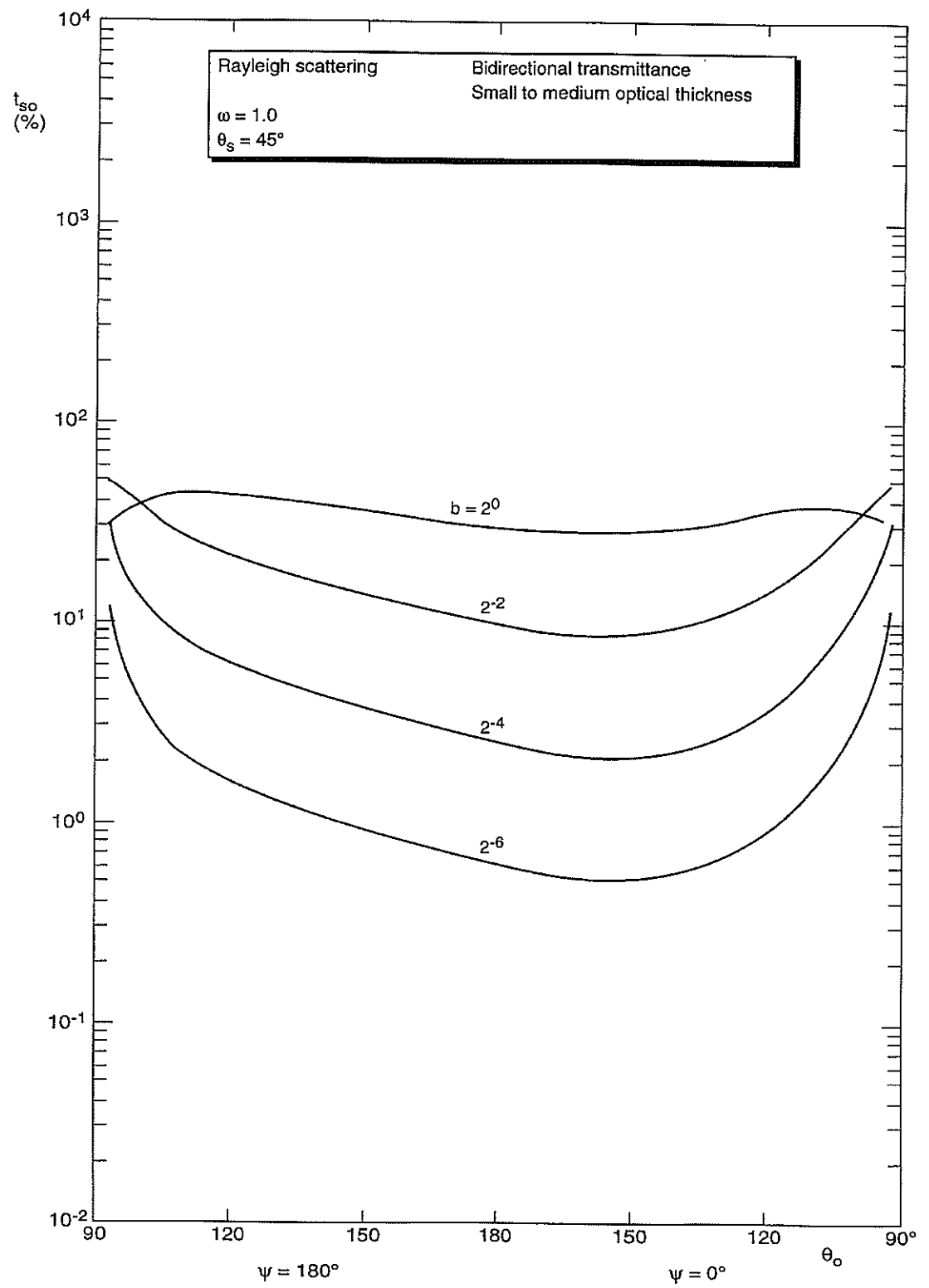


Fig. 6.16 Total bidirectional transmittance in the principal plane for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$. Small to medium optical thickness

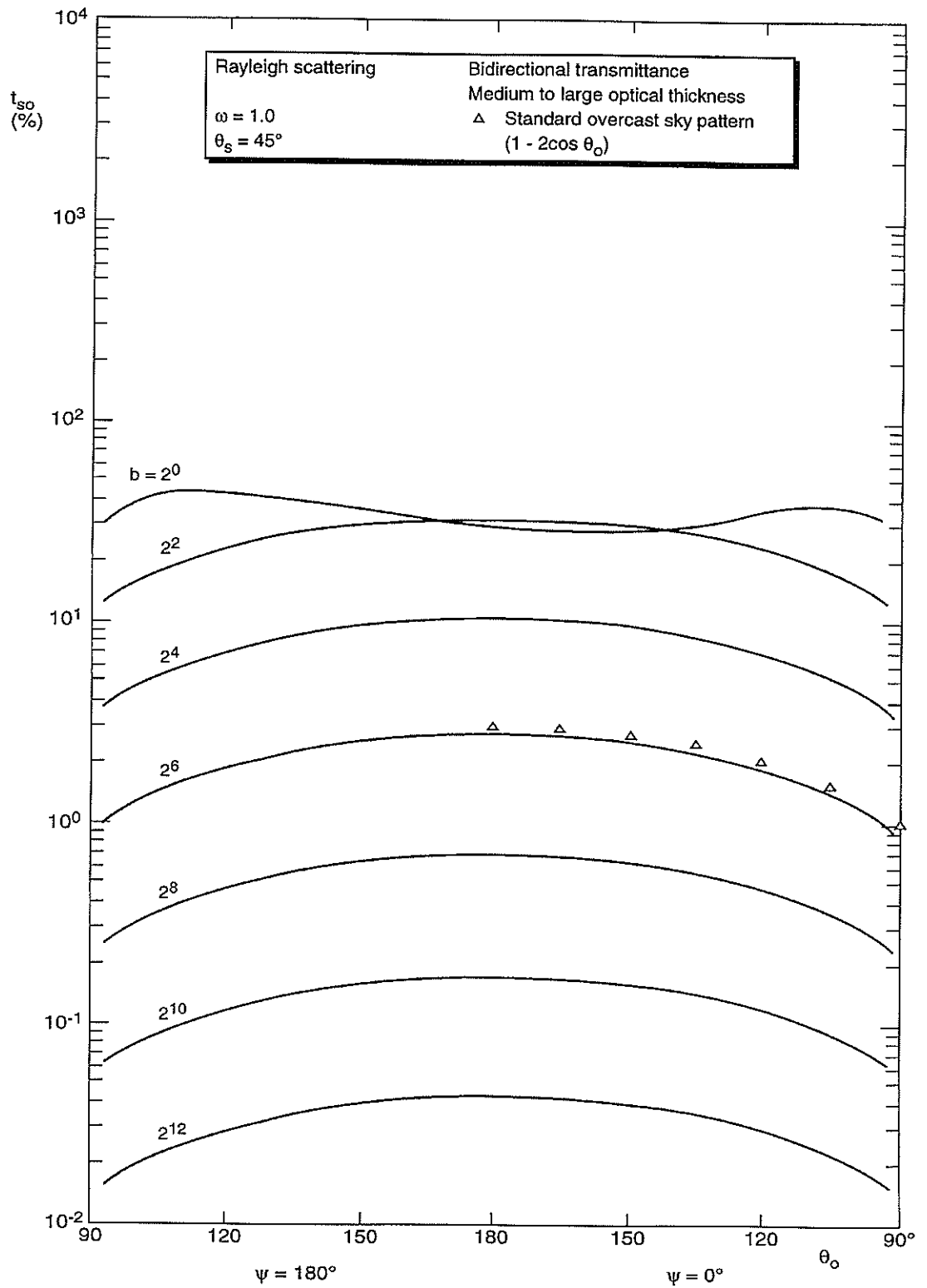


Fig. 6.17 Total bidirectional transmittance in the principal plane for Rayleigh scattering for $\omega = 1$ and $\theta_s = 45^\circ$. Medium to large optical thickness



pattern for the Standard Overcast Sky has been plotted in Fig. 6.17 too. Here it appears that this pattern is virtually equal to the model result, so it seems that this directional pattern indeed will result independently of the phase function. In Fig. 6.12 the (small) deviation from the Standard Overcast Sky pattern does not necessarily indicate a deficiency of the model, as asymptotic theory (Nakajima & King, 1992) only states that for optically thick layers the bidirectional transmittance is proportional to the product $K(\mu_s)K(\mu_o)$, where $K(\mu)$ is the so-called escape function. For the Standard Overcast Sky model this function happens to be given by $K(\mu) = 1 + 2\mu$, but this does not follow explicitly from the asymptotic theory, so there is room for deviation from this simple model.

Figures 6.18 to 6.21 demonstrate the reciprocity relations of the bidirectional reflectance and transmittance for $b = 2^{-4}$ and $b = 1$. For water Haze M the directional patterns are plotted for solar zenith angles θ_s of 15° , 45° and 75° . In these Figures points for the combination of angles $(\theta_s, \theta_o, \psi)$ have been connected by means of dashed lines with points for the combination of angles $(\theta_o, \theta_s, \psi)$. In order to fulfil the reciprocity relations, these lines should be strictly horizontal. In the Figures it appears they are, and also in the numerical results no differences at all have been found between corresponding pairs of simulation results. In view of chapter 5, where symmetric expressions for the bidirectional reflectance and transmittance were derived which automatically fulfil the reciprocity relations, it may not be surprising that the programmed calculations indeed give the expected results, so in this case the reciprocity relations have only been used as a means to check the correctness of programming.

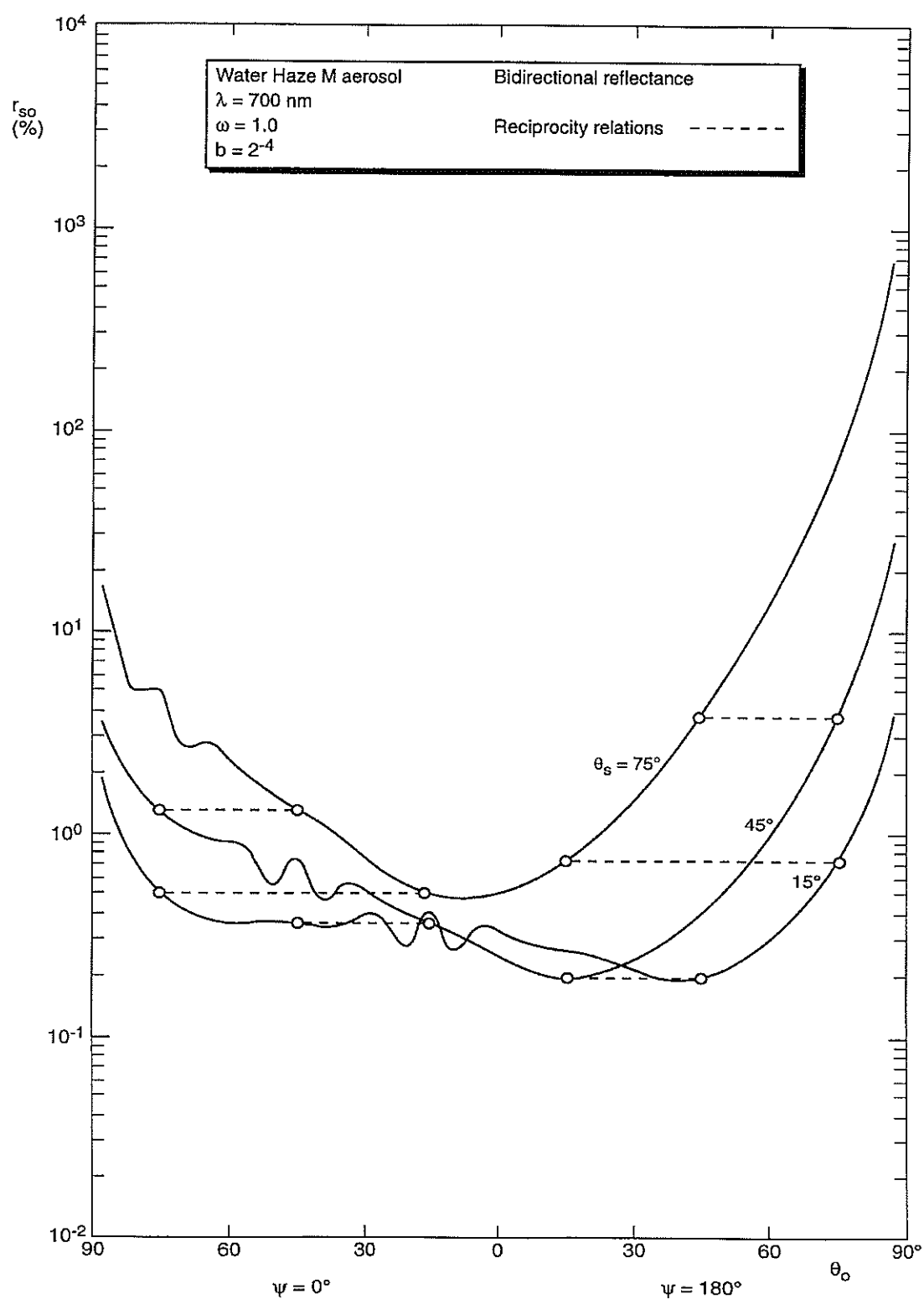


Fig. 6.18 Reciprocity relations of the bidirectional reflectance for water Haze M aerosol at 700 nm for $\omega = 1$ and $b = 2^{-4}$

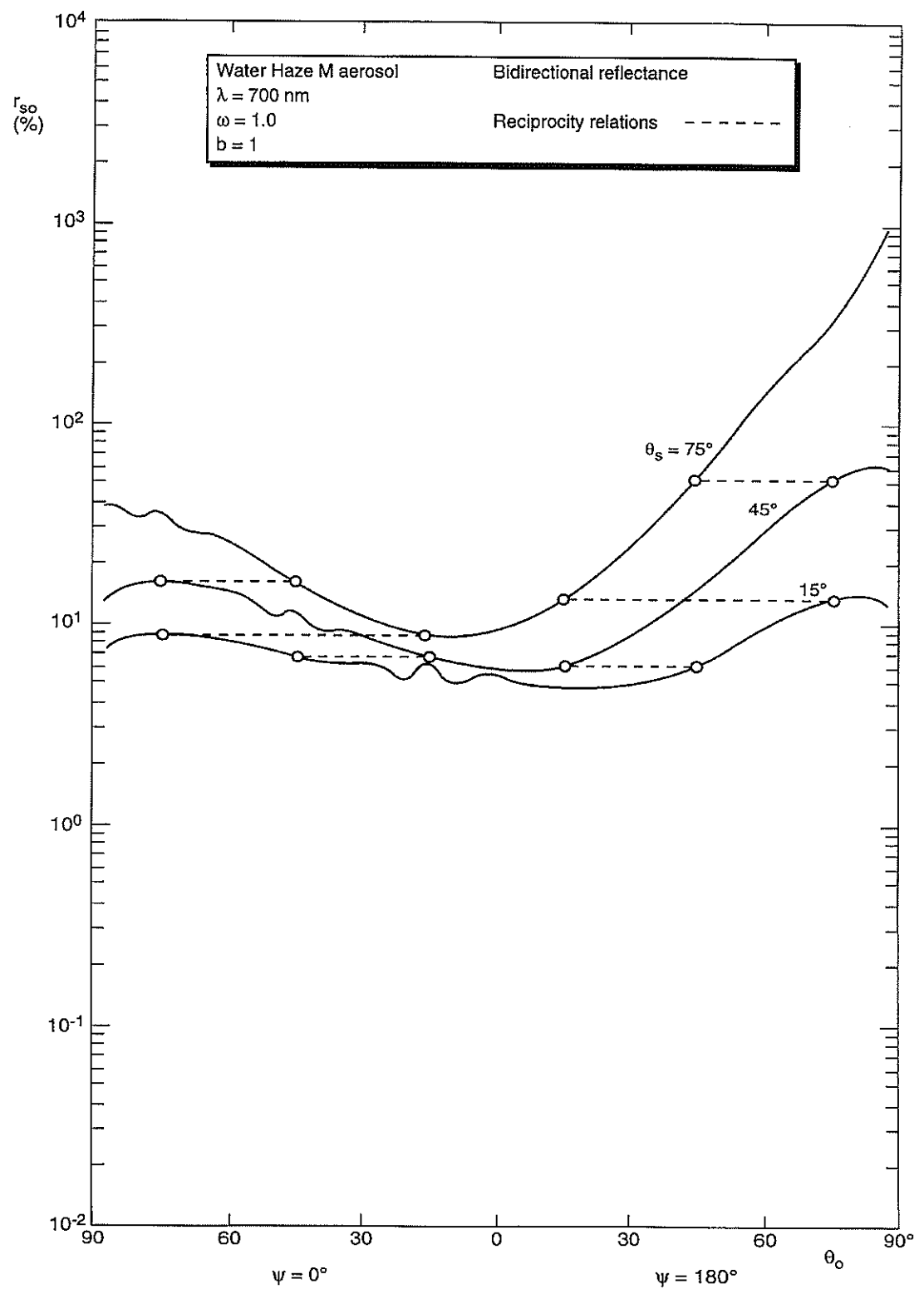


Fig. 6.19 Reciprocity relations of the bidirectional reflectance for water Haze M aerosol at 700 nm for $\omega = 1$ and $b = 1$

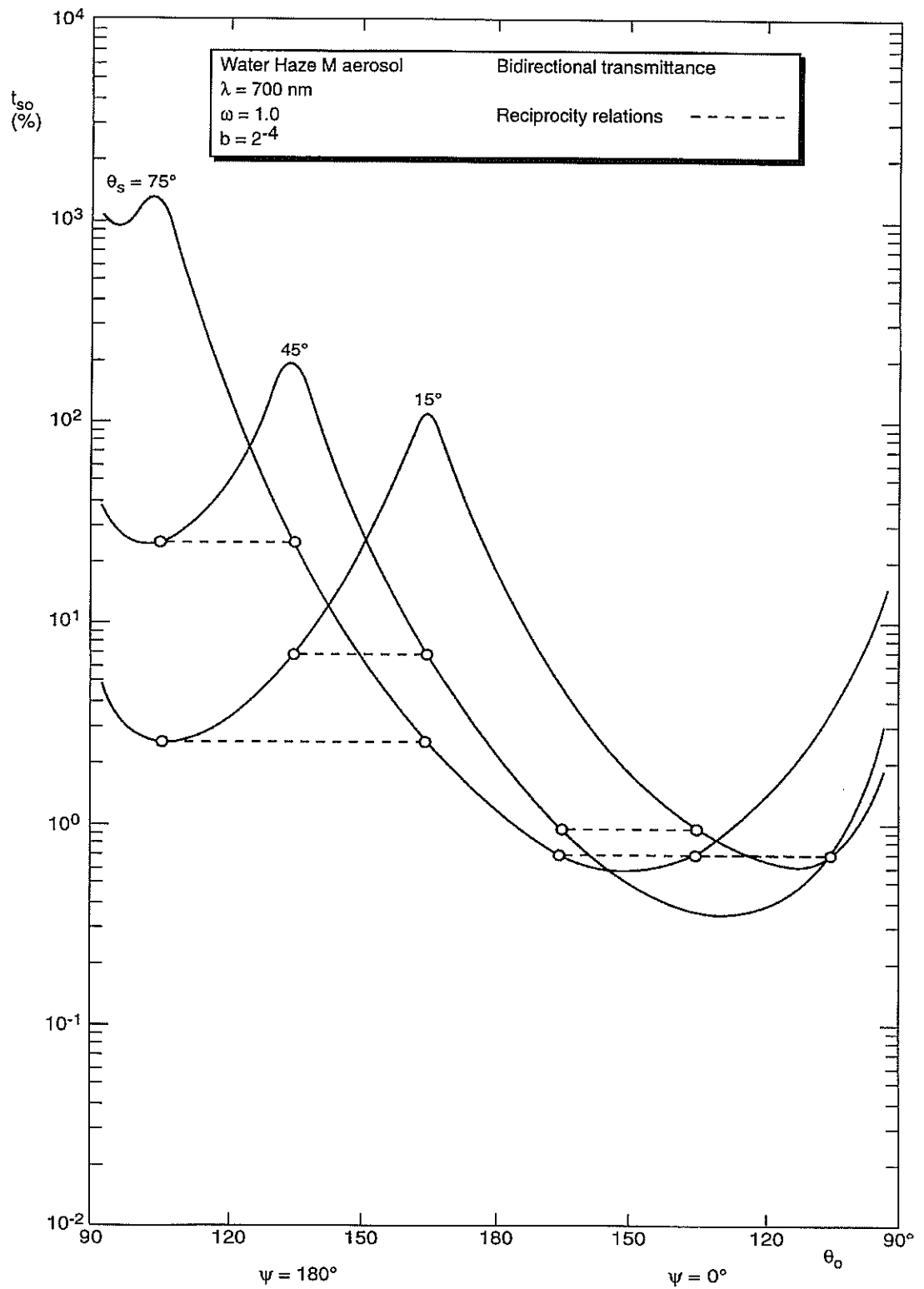


Fig. 6.20 Reciprocity relations of the bidirectional transmittance for water Haze M aerosol at 700 nm for $\omega = 1$ and $b = 2^{-4}$

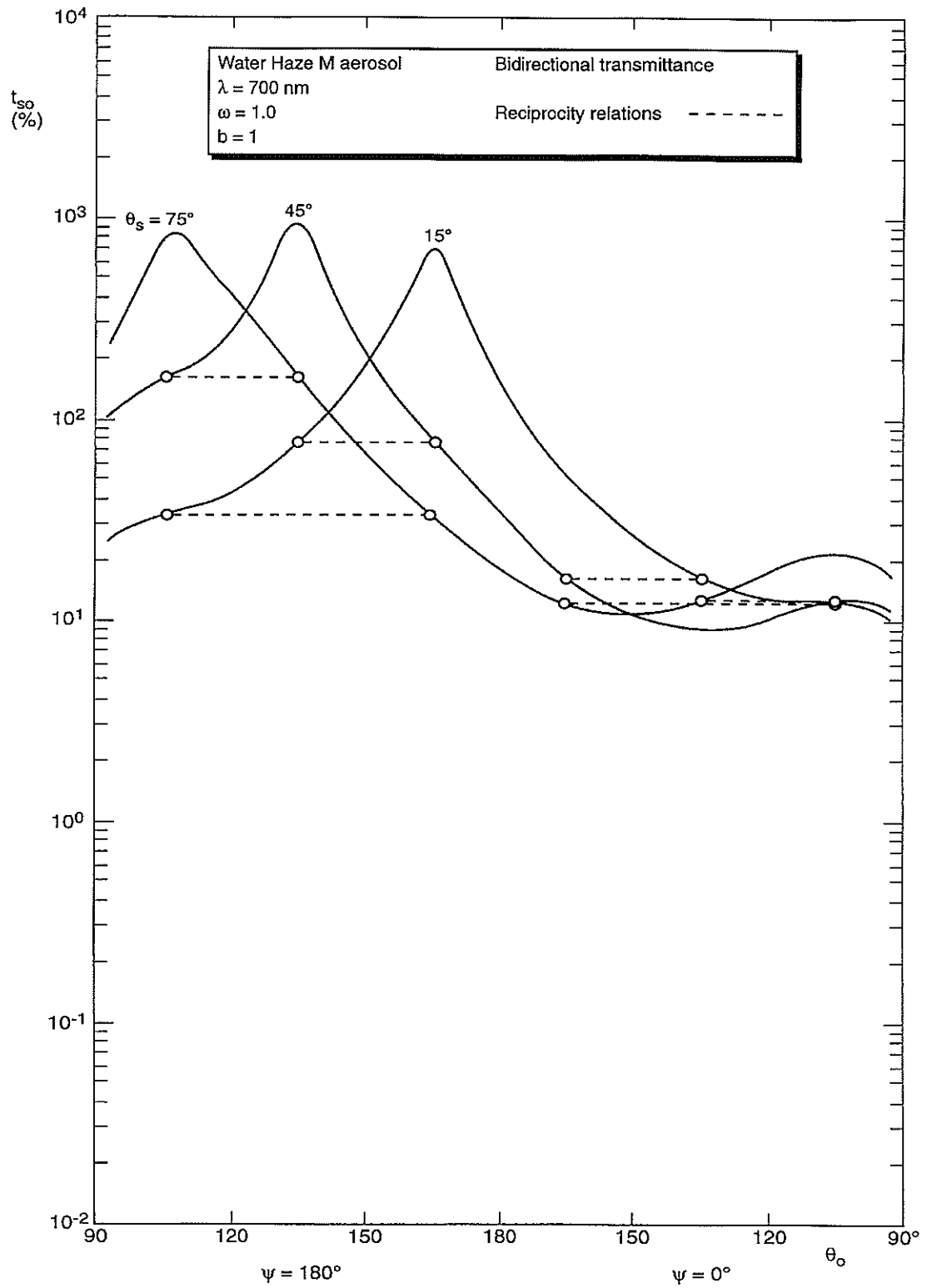


Fig. 6.21 Reciprocity relations of the bidirectional transmittance for water Haze M aerosol at 700 nm for $\omega = 1$ and $b = 1$



6.6 Discussion

The results presented in this chapter are not the first in which the performance of an $(N+2)$ -stream radiative transfer model is explored and compared with a four-stream model. In Verhoef (1988) already an attempt in this direction was made, but at that time no reliable simulations could be carried out because of numerical problems, which in particular came to expression as irregularities in directional profiles of the bidirectional transmittance. Since that time a lot of effort has been spent in locating the origins of these problems and the result is the entirely revised theory presented in chapter 5 and Appendix A. The old theory was a direct generalization of the four-stream theory presented in Verhoef (1985a), which resulted in a very compact formulation. The formulation of the new theory is less compact, but all potential numerical problems related with the possible equality of eigenvalues have been avoided. Yet, both theories are mathematically equivalent, which illustrates that the translation of a mathematical formulation into numerically safe computer code can be quite critical.

In the results of the previous sections no indication of any of those numerical problems has been found, in spite of the wide range of input conditions applied. The only indication of (another type of) numerical inaccuracy is the fact that for the aerosol phase function and conservative scattering the smallest computed eigenvalue was not zero but 0.000018. However, the numerically induced layer absorption caused by this is very small, even for optically very thick layers, so correction of this error, for instance by using double precision arithmetic, is considered not necessary.

Compared to the four-stream model, the $(N+2)$ -stream model allows a much more accurate calculation of the directional profile of the multiple scattering contribution to the bidirectional reflectance and transmittance of a layer. In the previous sections it has been shown that already at an optical thickness of 0.25 this contribution is of the same order of magnitude as the single scattering contribution, which implies that for the atmospheric conditions under which optical remote sensing is usually applied the four-stream model is inadequate. However, it is still more accurate than a single scattering approximation, which ignores multiple scattering completely, and for cases in which no high accuracy is required, or directional variations are of less concern, and for simulations at moderate zenith angles (maximum off-nadir angle 30°) the four-stream model might still be useful.

For the modelling of atmospheric radiative transfer the Discrete Ordinates



Method (DOM), which is based on the Gaussian quadrature method of angular integration, is well-established and probably more accurate (at comparable angular detail) than the $(N+2)$ -stream method, which is based on a simple integration over discrete angular segments. However, in most DOM-based models it is assumed that the medium is isotropic (i.e. its phase function and interception function are independent of the incidence direction), which means that they cannot be applied to vegetation canopies or other non-isotropic media. The $(N+2)$ -stream model presented in chapter 5 and Appendix A is only restricted to azimuthally isotropic media, which is a much broader category, as vegetation canopies with random leaf azimuth orientation also belong to it. Application of the $(N+2)$ -stream model to vegetation canopies is straightforward and can be based on the interception function and the volume scattering function as calculated in the SAIL model (cf. chapter 7). This would make it possible to obtain a better estimation of the multiple scattering contribution in vegetation canopies than obtained at present with the four-stream SAIL model. In most regions of the spectrum the absorption of radiation by leaves is strong, so that the benefits of an $(N+2)$ -stream version of SAIL would mainly be limited to the near infrared part of the spectrum, where absorption is weak, and multiple scattering plays a significant role. However, since the $(N+2)$ -stream model not only gives more accurate bidirectional reflectances but also a more refined description of the diffuse reflectance by means of vectors and matrices instead of scalar values, the light interaction between a vegetation canopy and the atmosphere can also be described more accurately. This is especially interesting with respect to the treatment of the aureole region around the sun: in the current four-stream interaction models the sky radiance is treated as a uniformly diffuse source of radiation input to a vegetation canopy, while a great part of the sky radiation comes from the aureole region and should be treated more like a directional source of radiation. In terms of an $(N+2)$ -stream interaction model the aureole region would be expressed by a few bright segments in the neighbourhood of the sun and the directional effects of this on the reflectance of the vegetation would automatically be incorporated.

The description of the SAIL model presented in the next chapter can be used as a basis for development of an $(N+2)$ -stream version. However, in chapter 8, which discusses the hot spot effect, it appears that the finite leaf size has an enormous impact on the single scattering contribution to the bidirectional reflectance, so it was considered more appropriate to include this effect in SAIL first.



7 LIGHT SCATTERING BY LEAF LAYERS: THE SAIL MODEL

7.1 Introduction

A physical description of the interaction of light with a vegetation canopy and the soil background can give insight in the relation between biophysical vegetation parameters and the bidirectional reflectance, which is an important quantity in the description of optical remote sensing systems. Essentially, these systems measure differences of the bidirectional reflectance of objects on the ground, and canopy reflectance modelling can help explain the measured differences from a physical point of view, thus supporting the interpretation of images.

From the review of canopy reflectance models in chapter 4 it may have become clear that the model developed by G.H. Suits (1972) has been very important for many investigations of canopy reflectance, both theoretically and by measurement in the field (Bunnik, 1978). Suits's model is the first instance of a four-stream radiative transfer model, an extension of the widely-known Kubelka-Munk theory (two-stream theory) with 1) specular incident flux from the sun and 2) radiance in the direction of view. These extensions are necessary for the calculation of the bidirectional reflectance.

In the Suits model the following assumptions are made:

- a canopy layer is horizontal, homogeneous and infinitely extended in the horizontal plane, i.e. the model is one-dimensional
- leaves are infinitesimal in size and can be replaced by their horizontal and vertical projections, of which the latter are random in azimuth
- leaves are bi-Lambertian (in reflectance and transmittance)
- the reflectance of the soil is Lambertian

The input parameters of Suits's model are:

- H = horizontally projected leaf area index
- V = vertically projected leaf area index
- ρ = single leaf reflectance
- τ = single leaf transmittance
- r_s = soil reflectance



θ_s = solar zenith angle
 θ_o = view zenith angle
 ψ = relative azimuth angle

An advantage of the Suits model is that the extinction and scattering coefficients of the radiative transfer equations are explicitly expressed in the above parameters (except r_s). This can be summarized by (cf. Bunnik, 1978)

$k = f(H, V, \theta_s)$
 $K = f(H, V, \theta_o)$
 $a = f(H, V, \rho, \tau)$
 $\sigma = f(H, V, \rho, \tau)$
 $s = f(H, V, \rho, \tau, \theta_s)$
 $s' = f(H, V, \rho, \tau, \theta_s)$
 $v = f(H, V, \rho, \tau, \theta_o)$
 $v' = f(H, V, \rho, \tau, \theta_o)$
 $w = f(H, V, \rho, \tau, \theta_s, \theta_o, \psi)$

The above extinction and scattering coefficients are introduced in chapter 4 (section 4.1.2). The non-Lambertian reflectance of a canopy is totally attributed to the vertical leaf projections, as a canopy consisting of solely horizontal leaves gives a Lambertian reflectance. In Suits's model the coefficients K , v , v' and w , which express the dependence of the bidirectional reflectance on the viewing direction, depend on the factor $V \tan \theta_o$. This factor generates V-shaped profiles of the bidirectional reflectance when plotted as a function of the viewing angle in the principal plane, which seems rather unrealistic. As this kind of behaviour is directly related with the vertical leaf projections, it was concluded that the rather drastic simplification of canopy morphology to just horizontal and vertical projections was the cause of this behaviour. In Verhoef (1984) it is demonstrated that much improvement can be made if this simplification is dropped, thus leading to the model SAIL (from Scattering by Arbitrarily Inclined Leaves). In Figures 7.1 and 7.2 several bidirectional canopy reflectance profiles in the green part of the spectrum at 550 nm are shown for the Suits and the SAIL model, respectively. These clearly illustrate the difference in behaviour of both models. The SAIL model considers leaves of arbitrary inclination under the assumption that the leaf azimuth distribution is still uniform. In the SAIL model the extinction and scattering coefficients are expressed in the leaf

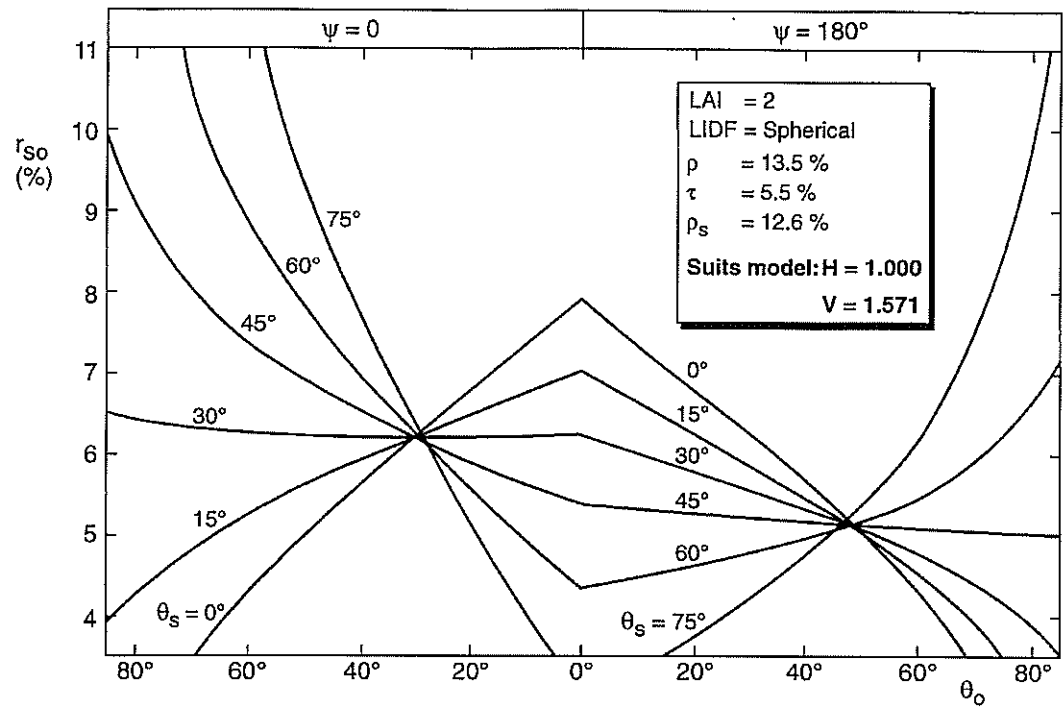


Fig. 7.1 Vegetation canopy bidirectional reflectance profiles in the green for the Suits model

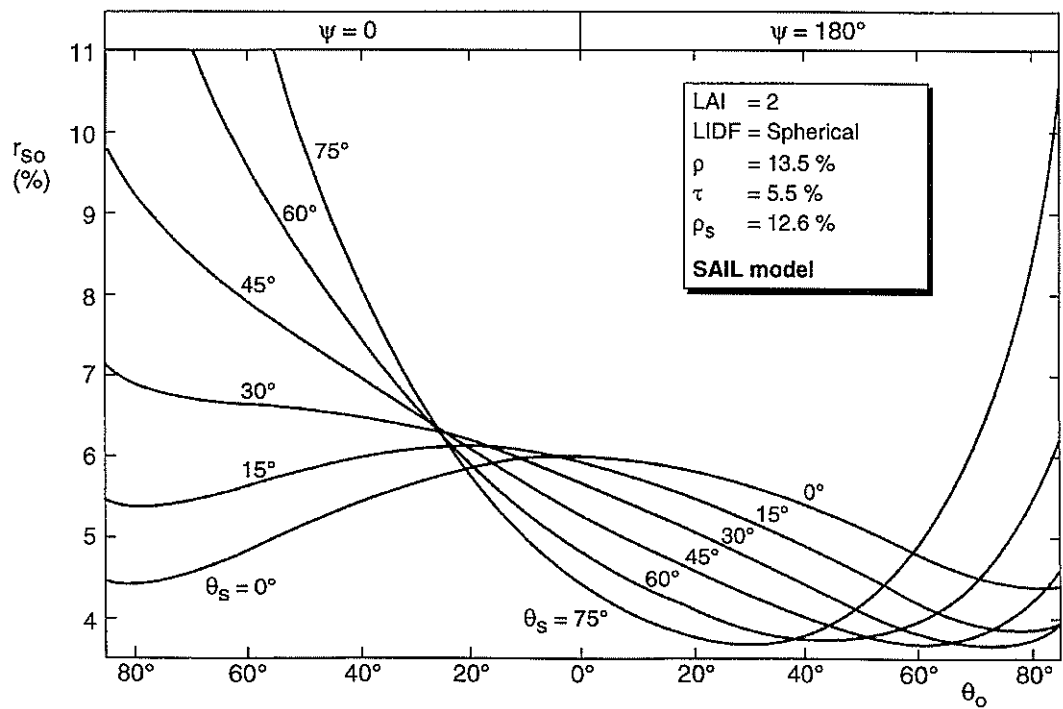


Fig. 7.2 Same as Fig. 7.1 but for SAIL model



inclination angle θ_i , and a leaf inclination distribution function (LIDF) is applied in order to form the weighted sum of these coefficients. Instead of the parameters H and V of Suits's model, the LIDF plus the total leaf area index (LAI) are used in the SAIL model. Within SAIL, the behaviour of Suits's model can still be imitated by specifying an LIDF consisting exclusively of fractions of horizontal and vertical leaf area, plus the total LAI. Therefore, it can be stated that the SAIL model includes the Suits model as a special case.

This chapter discusses the extinction and scattering coefficients of the SAIL model. The terminology has been adapted to that used elsewhere in this manuscript as far as possible. Examples of LIDFs are also discussed, and a method to describe these by means of only two parameters is proposed. Results of simulations with the SAIL model are presented in order to demonstrate the general behaviour of the model as a function of the viewing angle under various conditions.

At several institutes in the world the SAIL model has been tested against experimental data of crop reflectance in the field and remotely sensed data from aircraft and satellites. A detailed discussion of the results of these experiments falls outside the scope of this thesis, but the general conclusion may be that the model performs reasonably well for vegetation canopies which do not deviate too much from the assumptions on which the model is based (Goel, 1988). For instance, leaf gloss, row structure, clumpiness, and the finite leaf size are all factors not included in SAIL and one cannot expect to find good modelling results for crops which show any of these properties to a significant extent.

7.2 Extinction and scattering coefficients

The extinction and scattering coefficients of the SAIL model are the coefficients of the system of differential equations

$$\frac{d}{Ldx} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix} = \begin{bmatrix} k & & & \\ -s' & a & -\sigma & \\ s & \sigma & -a & \\ w & v & v' & -K \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix}, \quad (7.1)$$



where E_s = direct solar flux density on a horizontal plane
 E^- = diffuse downward irradiance
 E^+ = diffuse upward irradiance
 $E_o = \pi L_o$, where L_o is the radiance in the direction of view
 L = total leaf area index of the layer
 x = relative optical height within the layer.

The extinction coefficients are k (for direct solar flux E_s), K (for flux-equivalent radiance E_o) and κ (for the diffuse fluxes E^- and E^+). The scattering coefficients can be expressed by $\gamma(E_{in}, E_{out})$, where E_{in} denotes the type of incident flux and E_{out} the type of scattered flux, giving

$$\begin{aligned}
 \sigma &= \gamma(E^-, E^+) = \gamma(E^+, E^-) \\
 \sigma' &= \gamma(E^-, E^-) = \gamma(E^+, E^+) \\
 s &= \gamma(E_s, E^+) \\
 s' &= \gamma(E_s, E^-) \\
 v &= \gamma(E^-, E_o) \\
 v' &= \gamma(E^+, E_o) \\
 w &= \gamma(E_s, E_o)
 \end{aligned}$$

Note that κ and σ' do not appear as coefficients in Eq. (7.1). However, in Eq. (7.1) the diffuse attenuation coefficient a combines diffuse extinction and forward scattering via the equation $a = \kappa - \sigma'$, expressing that the net attenuation is the extinction caused by interception of diffuse flux minus the forward scattering of this flux.

The coefficients of the SAIL model depend on the morphological and optical properties of a leaf canopy layer, and also on the observational conditions. Thus, the parameters can be grouped as follows:

Morphological

L = leaf area index or LAI
 LIDF = leaf area inclination density function



Optical

ρ = single leaf reflectance
 τ = single leaf transmittance

Observational

θ_s = solar zenith angle
 θ_o = viewing zenith angle
 ψ = relative azimuth angle (viewing azimuth relative to solar azimuth)

Spherical coordinates are used to indicate directions of leaf normals, the sun and viewing by means of the unit vectors ℓ , s and o , respectively. The local normal vector is n . These vectors are illustrated in Fig. 7.3.

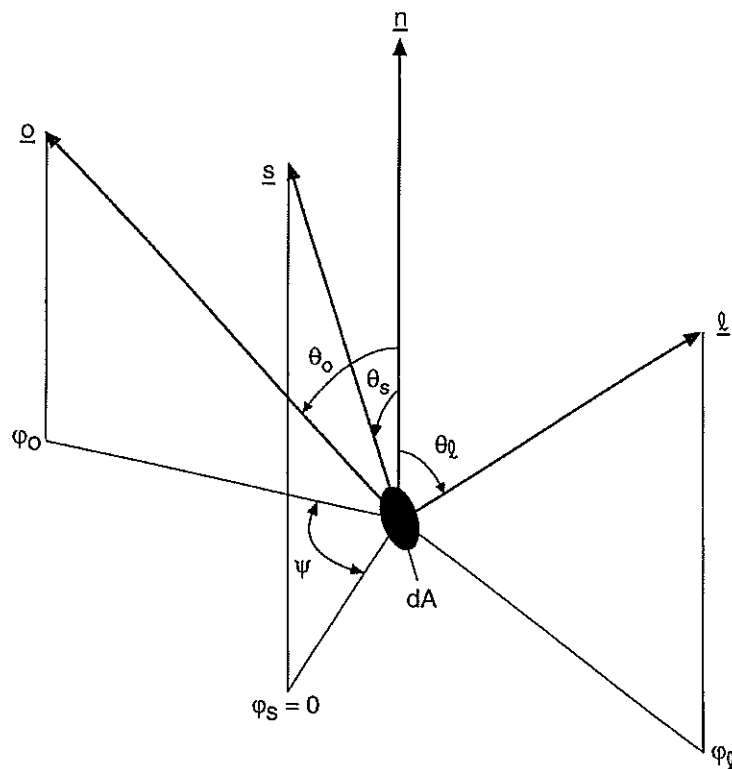


Fig. 7.3 Orientations of unit vectors ℓ , s and o relative to leaf area element dA and the normal vector n

Vector ℓ is the normal to the upper side of the leaf, so that it always points to the upper hemisphere. Vector s points to the sun and vector o points to



the observer. The vector normal to the layer in upward direction is called n . The angle ψ is given by $\psi = |\varphi_o - \varphi_s|$.

In the SAIL model it is assumed that the distribution of leaf azimuth is uniform. In that case only azimuthal differences are relevant, and without loss of generality one may set φ_s equal to zero.

For the determination of the coefficients it is first assumed that all leaves have identical orientation. In that case the transfer of radiation in a leaf layer is described by means of the projections of leaves and layer into the directions of the sun and observation, defined by the vector inner products

$$\cos \delta_s = \ell \cdot s \quad ; \quad \cos \theta_s = n \cdot s \quad ;$$

$$\cos \delta_o = \ell \cdot o \quad ; \quad \cos \theta_o = n \cdot o \quad .$$

The above projections become negative if the lower side of the leaves or the layer is illuminated or observed. In that case one has to take the absolute value of the projection if it is used in radiometric calculations.

When all the leaves have the same orientation, the SAIL coefficients can be based on the following three equations:

$$\text{I.} \quad d\Phi_i = A \cos \theta_s dE_s^o = E_s^o dA |\cos \delta_s|$$

$$\text{II.} \quad \pi \mathcal{E} = \begin{cases} \rho E_s^o \cos \delta_s & (\delta_s < \pi/2) \\ -\tau E_s^o \cos \delta_s & (\delta_s > \pi/2) \end{cases}$$

$$\text{III.} \quad dI_o = A |\cos \theta_o| dL_o = \mathcal{E} dA |\cos \delta_o|$$

The first equation describes the flux intercepted by a leaf layer when a source of specular incident radiation generates an irradiance E_s^o on a plane perpendicular to the incidence direction. The layer area considered is A and the leaf area fraction within an infinitesimal layer is dA . The absolute value of the projection of leaf area dA in direction s is taken in order to determine the intercepted flux (the right term), which is equivalent to the decrease of the irradiance dE_s^o multiplied by the projected layer area (middle term).

The second equation is used to determine the radiance $\mathcal{E}$ of the upper side of



the leaves induced by the irradiance E_s^o . For the lower side of the leaves one obtains the same equation, but with ρ and τ interchanged.

The third equation describes how the radiance of the leaves contributes to the increase of the radiance dL_o of the layer as a whole by stating that both are the cause of the same increase of the intensity (flux per unit solid angle) dI_o in the direction of observation. Note that for $\cos \theta_o$ also the absolute value is taken in order to ensure that positive contributions are obtained also in case this term becomes negative (which occurs for upward observations from inside or below the canopy layer, in which case $\theta_o > \pi/2$).

The leaf area index L of a canopy layer is the total one-sided leaf area per unit layer area. The relative optical height x within the layer ranges from -1 at the bottom to zero at the top, so the fraction of leaf area present in a vertical interval dx of the layer with area A equals $dA = LA dx$, so that $dA/A = L dx$. Substitution of this in equations I and III, and expressing the results in the fluxes $E_s = E_s^o \cos \theta_s$ and $E_o = \pi L_o$ gives

$$dE_s = E_s \frac{|\cos \delta_s|}{\cos \theta_s} L dx, \text{ and}$$

$$dE_o = \pi \mathcal{L} \frac{|\cos \delta_o|}{|\cos \theta_o|} L dx.$$

By means of the factors $f_s = \cos \delta_s / \cos \theta_s$ and $f_o = \cos \delta_o / |\cos \theta_o|$ one can write

$$dE_s = |f_s| E_s L dx \text{ and } dE_o = |f_o| \pi \mathcal{L} L dx.$$

Comparing the former equation with Eq. (7.1) learns that the extinction coefficient k for leaves of identical orientation equals $|f_s|$. For flux in the observer's direction the extinction coefficient $K = |f_o|$.

The bidirectional scattering coefficient γ is found from the latter equation by substitution of $\pi \mathcal{L}$ for upper and lower sides of leaves, which leads to the following table:



γ for:	$f_s > 0$	$f_s < 0$
$f_o > 0$	$f_s \rho f_o$	$-f_s \tau f_o$
$f_o < 0$	$-f_s \tau f_o$	$f_s \rho f_o$

From this it is seen that for a fixed leaf orientation the sign of the product $f_s f_o$ determines whether the interaction of incident specular flux with the leaves and the generation of radiance in the observer's direction is described by reflection or transmission.

From the above expressions, which apply only to specular fluxes, one can derive extinction and scattering coefficients for the case that the incident flux or the scattered flux, or both, are hemispherically diffuse (semi-isotropic), while the leaf orientation is still fixed. For this, use is made of the factors $f_1 = \frac{1}{2}(1 + \cos \theta_i)$ and $f_2 = \frac{1}{2}(1 - \cos \theta_i)$, which indicate the portions of diffuse hemispherical flux incident on or scattered from both sides of the leaf (see Appendix C). Summarized, this gives

Extinction coefficients:

$$\begin{aligned} k(\theta_i, \varphi_i) &= |f_s| \\ K(\theta_i, \varphi_i) &= |f_o| \\ \kappa(\theta_i, \varphi_i) &= f_1 + f_2 = 1 \end{aligned}$$

Scattering coefficients:

$$\begin{aligned} w(\theta_i, \varphi_i) &= f_s \rho f_o & \text{or} & \quad -f_s \tau f_o \\ s(\theta_i, \varphi_i) &= f_s (\rho f_1 + \tau f_2) & \text{or} & \quad -f_s (\tau f_1 + \rho f_2) \\ s'(\theta_i, \varphi_i) &= f_s (\tau f_1 + \rho f_2) & \text{or} & \quad -f_s (\rho f_1 + \tau f_2) \\ v(\theta_i, \varphi_i) &= (f_1 \rho + f_2 \tau) f_o & \text{or} & \quad -(f_1 \tau + f_2 \rho) f_o \\ v'(\theta_i, \varphi_i) &= (f_1 \tau + f_2 \rho) f_o & \text{or} & \quad -(f_1 \rho + f_2 \tau) f_o \\ \sigma(\theta_i, \varphi_i) &= f_1 (\rho f_1 + \tau f_2) + f_2 (\tau f_1 + \rho f_2) \\ \sigma'(\theta_i, \varphi_i) &= f_1 (\tau f_1 + \rho f_2) + f_2 (\rho f_1 + \tau f_2) \end{aligned}$$

The expressions on the right with the negative sign correspond to the cases for which $f_s f_o$, f_s , and f_o , respectively, is negative and are associated with:



- 1) illumination and observation of different sides of the leaf,
- 2) illumination of the lower side of the leaf,
- 3) observation of the lower side of the leaf.

For the extinction and scattering coefficients given above, the following relations, which demonstrate the conservation of radiant energy, apply:

$$\begin{aligned} s + s' &= (\rho + \tau) k \quad ; \\ v + v' &= (\rho + \tau) K \quad ; \\ \sigma + \sigma' &= (\rho + \tau) \kappa \quad . \end{aligned}$$

These equations express that of the intercepted flux a fraction $\rho + \tau$ is scattered over all directions. Since they apply for leaves of any orientation, they are also valid for any distribution of leaf orientations.

Fairly simple expressions for f_s and f_o are given by

$$\begin{aligned} f_s &= [C_s + S_s \cos \varphi_\ell] / \cos \theta_s \quad , \text{ and} \\ f_o &= [C_o + S_o \cos(\varphi_\ell - \psi)] / |\cos \theta_o| \quad , \\ \text{with } C_s &= \cos \theta_\ell \cos \theta_s \quad ; \quad S_s = \sin \theta_\ell \sin \theta_s \quad ; \\ C_o &= \cos \theta_\ell \cos \theta_o \quad ; \quad S_o = \sin \theta_\ell \sin \theta_o \quad . \end{aligned}$$

Considered as functions of the leaf azimuth angle φ_ℓ , the factors f_s and f_o can be seen to change sign at a certain leaf azimuth if $|S_s/C_s| > 1$ and $|S_o/C_o| > 1$, respectively. At this critical leaf azimuth illumination or observation of one side of the leaf changes to the other. The half leaf azimuth ranges for which the upper side of the leaf is illuminated and observed are called β_s and β_o . If there is no sign change, one may set $\beta_s = \pi$, and $\beta_o = \pi$ for $\theta_o < \pi/2$ and $\beta_o = 0$ for $\theta_o > \pi/2$.

The extinction and scattering coefficients for a distribution of leaf orientations are found by integrating those for fixed orientation given above over all possible orientations, with the leaf orientation density function $g(\theta_\ell, \varphi_\ell)$ as the weighting function.

In the SAIL model the integration over leaf azimuth φ_ℓ is carried out analytically under the assumption of a uniform leaf azimuth distribution. The



integration over leaf inclination θ_i is done numerically by summing the results obtained for uniform azimuth over a set of discrete leaf inclination angles, with a given leaf inclination distribution function (LIDF), which contains the frequencies of the discrete leaf inclination angles, as the weighting function.

The set of discrete leaf inclination angles consists of the series 5,15,...,75 and 81,83,...,89 degrees, representative for the intervals 0 - 10,...,70 - 80, and 80 - 82,...,88 - 90 degrees. The finer division of the interval 80 - 90 degrees is applied because for observation zenith angles close to zero, which is a situation found in many satellite remote sensing systems, some of the SAIL-coefficients are very sensitive to the LIDF in that range.

The analytical integrations carried out in order to find the azimuthal averages are simple for most of the coefficients, except for w . See Appendix C for detailed derivations. The results are summarized below. The extinction coefficients are given by

$$k(\theta_i) = \frac{2}{\pi} [(\beta_s - \pi/2) C_s + S_s \sin \beta_s] / \cos \theta_s, \quad ,$$

$$K(\theta_i) = \frac{2}{\pi} [(\beta_o - \pi/2) C_o + S_o \sin \beta_o] / \cos \theta_o, \quad , \text{ and}$$

$$\kappa(\theta_i) = 1 \quad ,$$

with for $C_s \geq S_s$: $\beta_s = \pi$, and for $C_s < S_s$: $\beta_s = \arccos(-C_s/S_s)$;

for $C_o \geq S_o$: $\beta_o = \pi$, and for $C_o < S_o$: $\beta_o = \arccos(-C_o/S_o)$.

The first six scattering coefficients are given by

$$\sigma(\theta_i) = \frac{1}{2} [\rho + \tau + (\rho - \tau) \cos^2 \theta_i] \quad ,$$

$$\sigma'(\theta_i) = \frac{1}{2} [\rho + \tau - (\rho - \tau) \cos^2 \theta_i] \quad ,$$

$$s(\theta_i) = \frac{1}{2} [(\rho + \tau) k(\theta_i) + (\rho - \tau) \cos^2 \theta_i] \quad ,$$

$$s'(\theta_i) = \frac{1}{2} [(\rho + \tau) k(\theta_i) - (\rho - \tau) \cos^2 \theta_i] \quad ,$$

$$v(\theta_i) = \frac{1}{2} [(\rho + \tau) K(\theta_i) + (\rho - \tau) \cos^2 \theta_i] \quad ,$$

$$v'(\theta_i) = \frac{1}{2} [(\rho + \tau) K(\theta_i) - (\rho - \tau) \cos^2 \theta_i] \quad .$$



The analytical integration over leaf azimuth for derivation of the bidirectional scattering coefficient $w(\theta_t)$ is rather difficult and is treated in Appendix C. The result is given by

$$w(\theta_t) = \frac{1}{2\pi} [\rho F_1(\theta_t) + \tau F_2(\theta_t)] \quad ,$$

$$\text{where } F_1(\theta_t) = \frac{(\pi - \beta_2)T_1(\theta_t) + T_2(\theta_t)}{|\cos\theta_s \cos\theta_o|} \quad ,$$

$$\text{and } F_2(\theta_t) = \frac{-\beta_2 T_1(\theta_t) + T_2(\theta_t)}{|\cos\theta_s \cos\theta_o|} \quad .$$

Here the functions $T_1(\theta_t)$ and $T_2(\theta_t)$ are given by

$$T_1(\theta_t) = 2 C_s C_o + S_s S_o \cos\psi \quad , \text{ and}$$

$$T_2(\theta_t) = \sin\beta_2 [2 D_s D_o + S_s S_o \cos\beta_1 \cos\beta_3] \quad ,$$

$$\text{where for } \beta_s < \pi : D_s = S_s \quad ,$$

$$\text{for } \beta_s = \pi : D_s = C_s \quad ,$$

$$\text{for } 0 < \beta_o < \pi : D_o = S_o \quad ,$$

$$\text{for } \beta_o = 0 \text{ or } \beta_o = \pi : D_o = -C_o / \cos\beta_o \quad .$$

The auxiliary azimuth angles β_1 , β_2 and β_3 are found from a decision table as follows:

Case	β_1	β_2	β_3
$\psi \leq \Delta_1$	ψ	Δ_1	Δ_2
$\Delta_1 < \psi < \Delta_2$	Δ_1	ψ	Δ_2
$\Delta_2 \leq \psi$	Δ_1	Δ_2	ψ



Finally, the angles Δ_1 and Δ_2 depend only on β_s and β_o by the simple expressions

$$\Delta_1 = |\beta_s - \beta_o| \quad \text{and} \quad \Delta_2 = \pi - |\beta_s + \beta_o - \pi| \quad .$$

With respect to the original article on the SAIL model (Verhoef, 1984) the following modifications have been introduced:

- The leaf area index L has been removed from the coefficients, because this is a factor common to all coefficients and because this better complies with the new notation of the differential equation (7.1).
- The coefficients s' , s , and u have been renamed to s , s' and v' , respectively. In this new notation all primed coefficients refer to forward scattering, whereas the unprimed refer to backscattering. Also, the association of s and s' with the sun and of v and v' with the viewing direction is better expressed now.
- The bidirectional scattering coefficient w has been completely reformulated in order to incorporate also the case $\theta_o > \pi/2$ (observation from below) and to better emphasize the fact that its angular dependence is fully described by the functions F_1 and F_2 , while the spectral dependence is given by ρ and τ as weighting factors for these functions.

The bidirectional scattering coefficient is one of the most important coefficients of the SAIL model, as it largely determines the so-called single scattering contribution to the canopy bidirectional reflectance. Especially in the visual spectral region this contribution is responsible for more than 90 percent of the total. Therefore, the behaviour of $w(\theta_i)$ will now be analyzed for a number of special cases.

I. Horizontal leaves ($\theta_i = 0$)

In this case $T_1(\theta_i) = 2 \cos \theta_s \cos \theta_o$, $\beta_s = \pi$, $\beta_o = \pi$ for $\theta_o < \pi/2$ and $\beta_o = 0$ for $\theta_o > \pi/2$. For the auxiliary azimuth angles one finds:



Case	Δ_1	Δ_2	β_1	β_2	β_3
$\theta_o < \pi/2$	0	0	0	0	ψ
$\theta_o > \pi/2$	π	π	ψ	π	π

In both cases $\sin \beta_2 = 0$, and therefore $T_2(\theta_t) = 0$ also. The expressions for F_1 and F_2 then give

$$\begin{aligned} F_1 &= 2\pi , \quad F_2 = 0 \quad \text{for } \theta_o < \pi/2 , \text{ and} \\ F_1 &= 0 , \quad F_2 = 2\pi \quad \text{for } \theta_o > \pi/2 . \end{aligned}$$

This means that one obtains $w = \rho$ for $\theta_o < \pi/2$ and $w = \tau$ for $\theta_o > \pi/2$, as could be expected.

II. Vertical leaves ($\theta_t = \pi/2$)

In this case $T_1(\theta_t) = \sin \theta_s \sin \theta_o \cos \psi$, $\beta_s = \pi/2$ and $\beta_o = \pi/2$, which gives $\Delta_1 = 0$, $\Delta_2 = \pi$, so that $\beta_1 = 0$, $\beta_2 = \psi$ and $\beta_3 = \pi$, which leads to $T_2(\theta_t) = \sin \psi (2 \sin \theta_s \sin \theta_o - \sin \theta_s \sin \theta_o) = \sin \theta_s \sin \theta_o \sin \psi$,

so that $F_1 = [(\pi - \psi) \cos \psi + \sin \psi] |\tan \theta_s \tan \theta_o|$, and

$$F_2 = [-\psi \cos \psi + \sin \psi] |\tan \theta_s \tan \theta_o| .$$

Substitution in the formula for $w(\theta_t)$ then gives

$$w = \frac{1}{2\pi} \{ \rho [(\pi - \psi) \cos \psi + \sin \psi] + \tau [-\psi \cos \psi + \sin \psi] \} |\tan \theta_s \tan \theta_o| .$$

This expression is entirely consistent with the corresponding one published in Verhoef (1984).

III. The hot spot ($\theta_o = \theta_s$, $\psi = 0$)

In this case the direction of view coincides exactly with the direction of the sunrays, so that $\beta_o = \beta_s$ and $\Delta_1 = 0$. Since also $\psi = 0$, it follows



that $\beta_1 = \beta_2 = 0$, and therefore $T_2(\theta_t) = 0$. For $T_1(\theta_t)$ one finds

$$T_1(\theta_t) = 2 \cos^2\theta_t \cos^2\theta_s + \sin^2\theta_t \sin^2\theta_s, \text{ so that}$$

$$F_1 = \pi (2 \cos^2\theta_t + \sin^2\theta_t \tan^2\theta_s) \text{ and } F_2 = 0.$$

For $w(\theta_t)$ this gives

$$w(\theta_t) = \rho (\cos^2\theta_t + \frac{1}{2} \sin^2\theta_t \tan^2\theta_s).$$

It is interesting to note that this expression becomes insensitive to the leaf inclination angle θ_t if $\tan^2\theta_s = 2$, or $\theta_s = \arctan(\sqrt{2}) = 54.7$ degrees. For a detailed discussion on more of these "special" angles, associated with the extinction coefficient k , the bidirectional scattering coefficient w , and the hot spot reflectance, all for the Suits model, the reader is referred to Bunnik (1978). For the SAIL model the sensitivity of the hot spot reflectance to the leaf inclination does not become zero but rather reaches a minimum in the range of θ_s between 50 and 60 degrees.

In addition to the properties discussed above, some more general characteristics of the bidirectional scattering coefficient will be discussed now.

It obeys the reciprocity relation, which means that $w(\theta_t)$ remains invariant when the directions of the sun and of observation are interchanged. The proof is trivial, since from the equations forming $w(\theta_t)$ it can be concluded that both $T_1(\theta_t)$ and $T_2(\theta_t)$ are completely symmetric with respect to the quantities related with the two directions, so this also holds for $F_1(\theta_t)$ and $F_2(\theta_t)$, and therefore also for $w(\theta_t)$. Mathematically, this property can be expressed by $w(\theta_s, \theta_o, \psi) = w(\theta_o, \theta_s, \psi)$.

Another property of $w(\theta_t)$ is based on the functions $F_1(\theta_t)$ and $F_2(\theta_t)$. These functions interchange when the direction of observation is completely reversed (θ becomes $-\theta$). One may express this by

$$F_1(\theta_o, \psi) = F_2(\pi - \theta_o, \pi - \psi) \quad , \text{ and}$$

$$F_2(\theta_o, \psi) = F_1(\pi - \theta_o, \pi - \psi) \quad .$$



The effect on $w(\theta_i)$ is that by this operation the roles of ρ and τ interchange, or

$$w(\rho, \tau, \theta_o, \psi) = w(\tau, \rho, \pi - \theta_o, \pi - \psi) \quad .$$

The proof of this property is based on the change of sign of $T_1(\theta_i)$ and the replacement of the term $\pi - \beta_2$ by β_2 . The term $T_2(\theta_i)$ remains invariant under this operation.

Next follows a discussion of the two basic quantities describing the transfer of radiation in a leaf canopy, namely the interception coefficient and the volume scattering coefficient. These quantities are important because they do not become infinite for horizontal directions of incidence or observation, and are better suited for incorporation in an (N+2)-stream model, as discussed in chapter 5.

In chapter 2 it was shown that $k = \beta / \mu_s$, where k is the extinction coefficient (for specular flux), β is the interception coefficient and $\mu_s = \cos \theta_s$. From this it follows that $\beta = k \cos \theta_s$ and for a leaf inclination angle θ_i one finds

$$\beta(\theta_i) = \frac{2}{\pi} [(\beta_s - \pi/2) C_s + S_s \sin \beta_s] \quad .$$

A special case is formed by the so-called spherical leaf angle distribution. In that case the interception coefficient is the average absolute projection in the direction of the sun, which is equal to 0.5, independent of the solar incidence angle. However, for any other leaf angle distribution the interception coefficient is a function of the incidence angle.

Examples of the interception coefficient $\beta(\theta_i)$ as a function of the solar zenith angle θ_s are shown in Fig. 7.4. In this case the curve for horizontal leaves ($\theta_i = 0$) is simply $\cos \theta_s$ and that for vertical leaves ($\theta_i = \pi/2$) equals $(2/\pi) \sin \theta_s$. They intersect at $\theta_s = \arctan(2/\pi) = 57.5$ degrees, another "magic" angle (Bunnik, 1978) in relation with the Suits model, since at this angle the extinction coefficient k is independent of the proportions of horizontal and vertical leaf area. The incorporated curves for $\theta_i = 30^\circ$ and $\theta_i = 60^\circ$ show that according to the SAIL model the sensitivity to the leaf inclination cannot become zero any more, but still a minimum is reached between 50° and 60° solar zenith angle.

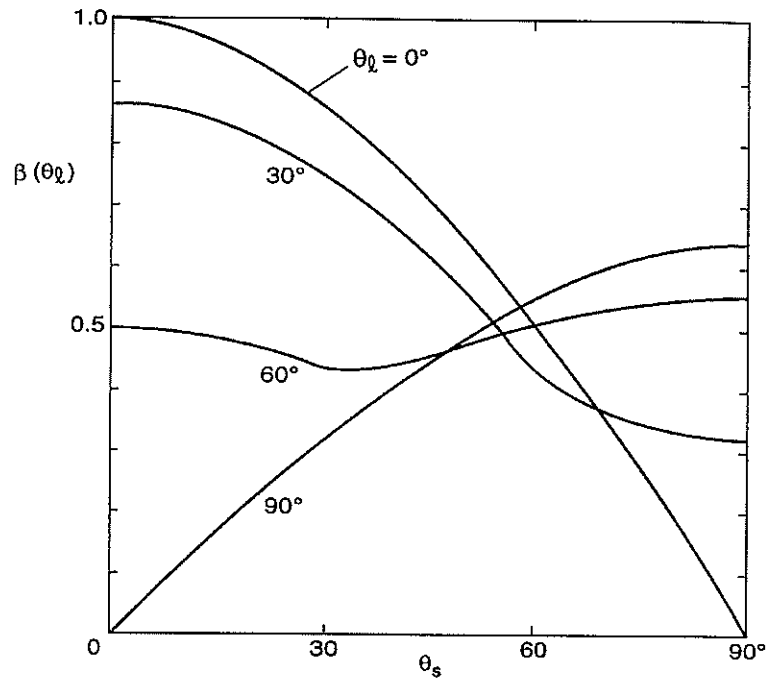


Fig. 7.4 The interception coefficient as a function of θ_s for four values of the leaf inclination angle θ_l

According to chapter 2, the relation between the bidirectional scattering coefficient (or area scattering coefficient) γ and the volume scattering function γ' is given by $\gamma' = \gamma \mu_s \mu_o / \pi$.

In the SAIL model $w(\theta_l)$ is the bidirectional scattering coefficient, so that the volume scattering function for the SAIL model is the quantity $\gamma'(\theta_l)$, defined by $\gamma'(\theta_l) = w(\theta_l) |\cos\theta_s \cos\theta_o| / \pi$, and is given by

$$\gamma'(\theta_l) = \frac{1}{2\pi^2} \{ \pi \rho T_1(\theta_l) + (\rho + \tau) [-\beta_2 T_1(\theta_l) + T_2(\theta_l)] \} .$$

For the same leaf inclination angles as in Fig. 7.4 the volume scattering function $\gamma'(\theta_l)$ is shown in Fig. 7.5 as a function of the observation angle θ_o in the principal plane for a solar zenith angle of 35° . The ratio $\rho / (\rho + \tau)$ equals 0.7, which is representative for the visible region of the spectrum. In this case the scattering in backward directions is stronger than in forward directions. In order to bring out this difference more clearly, a complete "circular" scan is shown in Fig. 7.5.

Like was the case also for the interception coefficient $\beta(\theta_l)$, the volume scattering functions for horizontal and vertical leaves are the absolute values

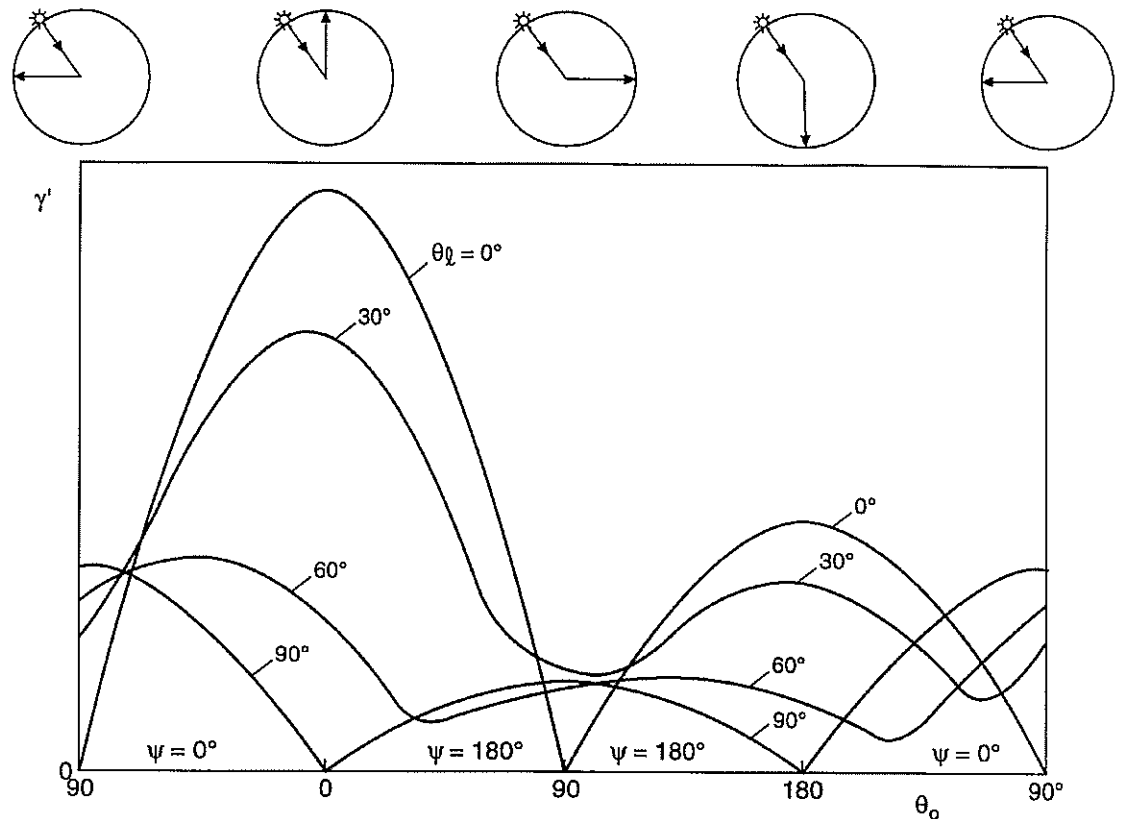


Fig. 7.5 Examples of the volume scattering function on arbitrary scale as a function of θ_o for $\theta_s = 35^\circ$ in the principal plane

of simple cosine and sine functions, of which the relative amplitudes are mainly determined by the ratio $\rho / (\rho + \tau)$ and by the solar zenith angle. The curves for $\theta_l = 30^\circ$ and $\theta_l = 60^\circ$ are smooth over the complete scan, i.e. the first derivative is continuous for all θ_o , and cannot be reproduced as linear combinations of the curves for horizontal and vertical leaves, which in fact is done in Suits's model. The curves for horizontal and vertical leaves are very special, since they are the only ones having discontinuous first derivatives. In a real leaf canopy there is always a distribution of leaf inclinations, and horizontal and vertical leaves will occur with equally low frequency as any other exact inclination angle, namely zero. Therefore, for real canopies the volume scattering function and the interception coefficient will always be smooth functions of the angles θ_s and θ_o .



7.3 Leaf inclination distribution functions

For the study of photosynthesis in leaf canopies, De Wit (1965) introduced a classification of leaf inclination distributions in four major types, namely

planophile	: horizontal leaves are most frequent
erectophile	: vertical leaves are most frequent
plagiophile	: oblique leaves are most frequent
extremophile	: oblique leaves are least frequent

Other well-known distributions are the uniform distribution (all inclinations are equally frequent) and the spherical distribution (identical to the inclination distribution of the surface elements of a sphere).

In this section the leaf inclination density function is denoted as $f(\theta_i)$ and the cumulative leaf inclination distribution as $F(\theta_i)$, such that

$$dF(\theta_i)/d\theta_i = f(\theta_i).$$

For the unique identification of distribution functions it is useful to characterize them by means of analytical expressions. In this way the uniform distribution is specified by $f(\theta_i) = 2/\pi$ and $F(\theta_i) = (2/\pi)\theta_i$. For the spherical distribution $f(\theta_i) = \sin \theta_i$ and $F(\theta_i) = 1 - \cos \theta_i$.

In Verhoef & Bunnik (1975) and Bunnik (1978) analytical prototypes for the distributions introduced by De Wit were constructed by means of trigonometric functions as follows:

$$f(\theta_i) = \frac{2}{\pi} [1 + \cos(2\theta_i)] \quad (\text{planophile}) ,$$

$$f(\theta_i) = \frac{2}{\pi} [1 - \cos(2\theta_i)] \quad (\text{erectophile}) ,$$

$$f(\theta_i) = \frac{2}{\pi} [1 - \cos(4\theta_i)] \quad (\text{plagiophile}) ,$$

$$f(\theta_i) = \frac{2}{\pi} [1 + \cos(4\theta_i)] \quad (\text{extremophile}) .$$

Some authors prefer to use beta-functions (Goel & Strebel, 1984) to characterize leaf inclination distributions. An advantage of this approach is that the distribution can be specified by means of two statistical measures, namely



the mean leaf inclination and the standard deviation. However, bimodal distributions such as the extremophile distribution are poorly represented by these functions.

For the SAIL model (Verhoef & Bunnik, 1981) a number of leaf inclination distribution functions were generated by means of a "graphical" method. This method takes as a basis the cumulative distribution for the uniform case (Fig. 7.6).

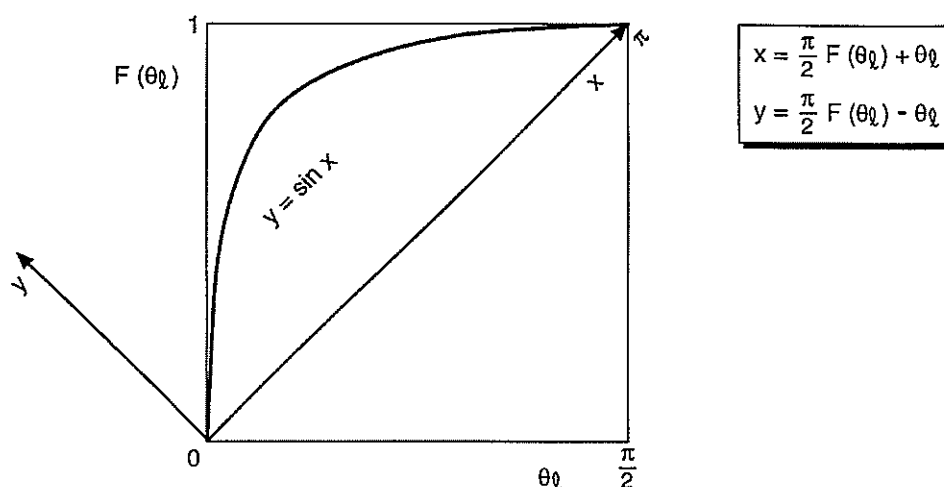


Fig. 7.6 Illustration of graphical method and co-ordinate transformation

In this diagram the diagonal represents the uniform distribution, and this diagonal is now employed as the X-ordinate for the trigonometric functions $a \sin x$ and $\frac{1}{2}b \sin 2x$. The Y-ordinate represents the deviation from the uniform distribution. As an example, Fig. 7.6 shows how the function $y = \sin x$ can be used as a prototype for the planophile distribution. Other distributions can be obtained by taking other combinations of a and b . The only restriction is that these parameters must be chosen such that $|a| + |b| < 1$, since otherwise the requirement that $F(\theta_l)$ must be a monotonously increasing function would be violated.

Parameter a controls the average leaf inclination, whereas parameter b influences the shape of the distribution (bimodality), but has no effect on the average leaf inclination.

For a given value of θ_l , the corresponding $F(\theta_l)$ can be found by means of a fast iteration. The pseudo code for this procedure is as follows:

$x = 2 \theta_t$
 Repeat
 $y = a \sin x + \frac{1}{2} b \sin 2x$
 $\Delta x = \frac{1}{2} (y - x + 2 \theta_t)$
 $x = x + \Delta x$
 Until $|\Delta x| < t$
 $F(\theta_t) = 2 (y + \theta_t) / \pi$

Here t is a threshold which should be set at a small value, e.g. 10^{-6} .

The average leaf inclination angle $\bar{\theta}_t$ is found as

$$\bar{\theta}_t = \int_0^{\pi/2} f(\theta_t) \theta_t d\theta_t = \int_0^{\pi/2} \frac{dF(\theta_t)}{d\theta_t} \theta_t d\theta_t = \int_0^1 \theta_t dF(\theta_t) = \frac{1}{2\pi} \int_0^1 (x-y) d(x+y) \quad .$$

Substitution of y and setting $d(x+y) = (1 + \frac{dy}{dx}) dx$ gives the solution

$$\bar{\theta}_t = \frac{\pi}{4} - \frac{2}{\pi} a \quad \text{radians} \quad .$$

$$\text{In degrees this gives } \bar{\theta}_t = 45^\circ - \frac{360}{\pi^2} a \quad .$$

For $a = 1$ one obtains $\bar{\theta}_t = 8.52^\circ$ and for $a = -1$ $\bar{\theta}_t = 81.48^\circ$. This range is much larger than the one found for the planophile and erectophile distributions in Bunnik (1978), which have $\bar{\theta}_t = 26.8^\circ$ and $\bar{\theta}_t = 63.2^\circ$, respectively.

In Fig. 7.7 it is demonstrated that by choosing proper combinations of a and b many different cumulative distributions can be generated, and that De Wit's types are very well represented by the distributions numbered 1, 9, 4 and 6. These are also the most extreme ones. More intermediate cases are given by the distributions numbered 2, 3, 7 and 8. Number 5 represents the uniform distribution.

Unfortunately, the spherical distribution cannot be represented in this manner. However, a reasonably good approximation can be achieved by the combination of parameters $a = -0.35$ and $b = -0.15$. In Fig. 7.7 on the right this point in the parameter space is marked by a cross.

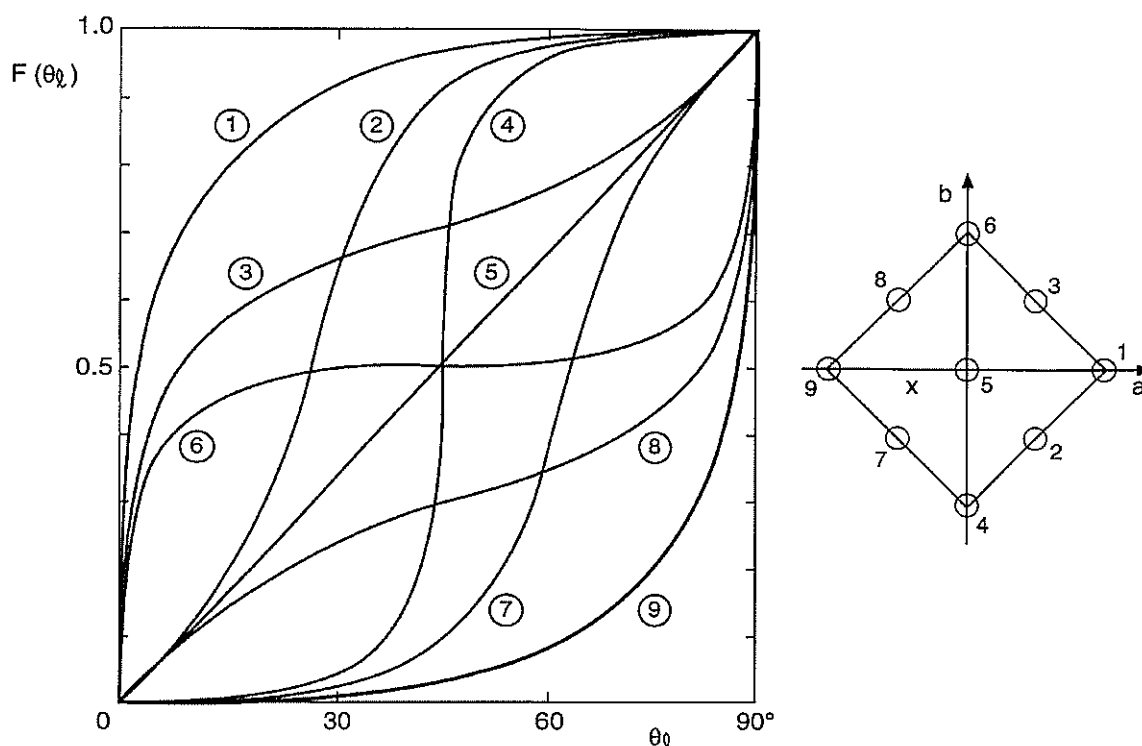


Fig. 7.7 Cumulative leaf inclination distribution functions generated by graphical method. Distributions in (a,b) parameter space on the right

The SAIL model uses discretized versions of distributions 1 to 9 and the spherical distribution (numbered 10). Discretization is carried out by calculation of $F(\theta_l)$ at the boundaries of the intervals and assigning the differences to the central leaf inclination of each interval. Table 7.1 gives the discrete leaf inclination distribution functions (LIDFs) so obtained.

The effects of leaf inclination distributions on the interception coefficient β and the volume scattering function γ' are demonstrated in Figures 7.8 and 7.9, respectively. Direct comparison with Figures 7.5 and 7.6 of the previous section, which showed these quantities for particular leaf inclination angles, is possible since the other parameters are identical in corresponding cases.

Because the leaf inclination distributions contain no exactly horizontal or vertical leaves (the closest are 5° and 89° inclination), the curves have become smoother. The curves for LIDF no. 1 (planophile) and no. 9 (erectophile) still resemble those for $\theta_l = 0^\circ$ and $\theta_l = 90^\circ$, respectively.

LIDFs 4 and 6 both have an average leaf inclination of 45° , but the former

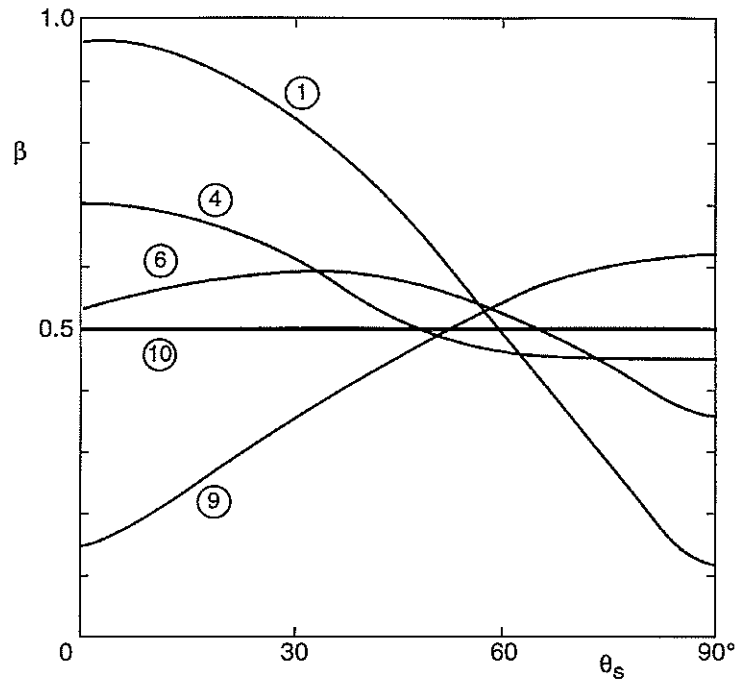


Fig. 7.8 The interception coefficient β as a function of θ_s for LIDF numbers 1, 4, 6, 9 and 10

is plagiophile and the latter is extremophile. It is interesting to note that in Figures 7.8 and 7.9 this difference in distribution type is clearly expressed: LIDF 6 behaves like a weighted average of LIDFs 1 and 9, whereas LIDF 4 (most leaves 45° inclination) differs from that by its own characteristic patterns of radiation interception and scattering.

For LIDF no. 10 (spherical) the interception coefficient is constant and equal to 0.5, which is confirmed in Fig. 7.8. In the numerical results for LIDF no. 10 the interception coefficient varied between 0.499 and 0.502, which indicates a good approximation of the theoretical value. This means that the discretization applied introduces only minor distortions.

The curve for the spherical distribution no. 10 in Fig. 7.9 shows that, contrary to the interception coefficient, its volume scattering function is not constant. However, for this example of a so-called isotropic medium the angular behaviour of the volume scattering function is relatively simple, as it depends only on the scattering angle δ_{so} , which can be found from $\cos \delta_{so} = s \cdot o$. The mathematical expression for this dependence is given by

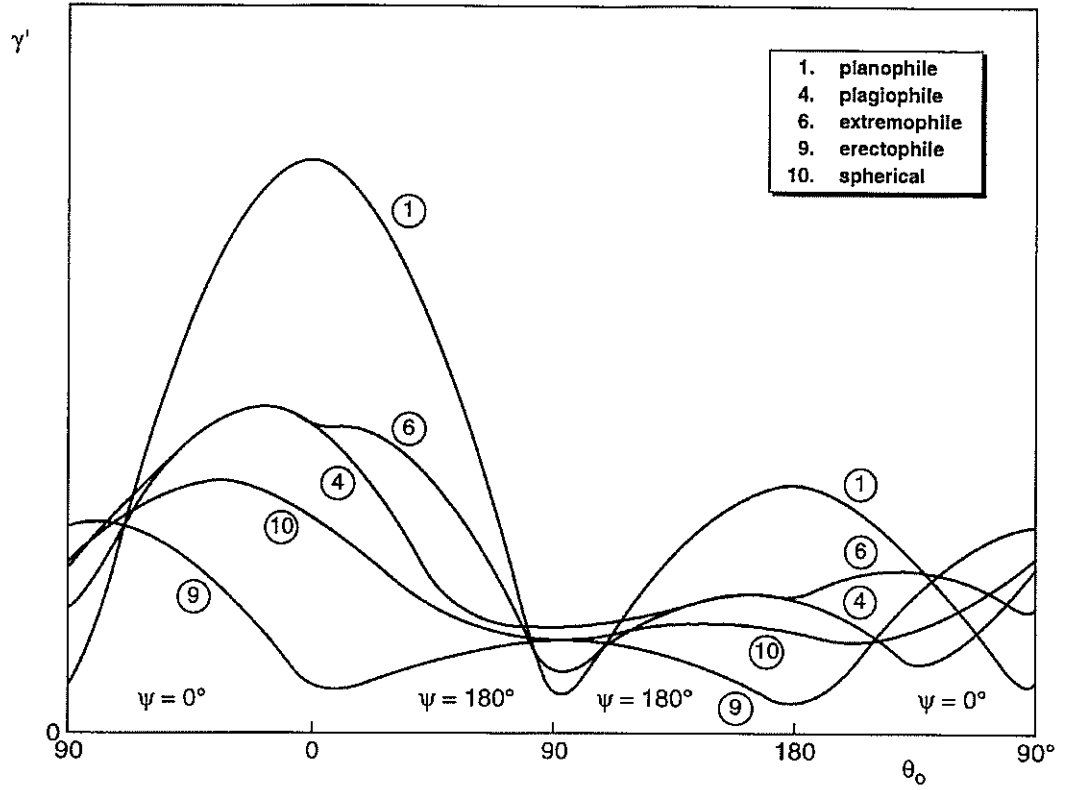


Fig. 7.9 Volume scattering functions as a function of θ_o for $\theta_s = 35^\circ$ in the principal plane and LIDF numbers 1, 4, 6, 9 and 10

$$\gamma'(\text{spherical}) = \frac{1}{3\pi^2} [\pi\rho \cos\delta_{so} + (\rho + \tau)(-\delta_{so} \cos\delta_{so} + \sin\delta_{so})] \quad .$$

This fairly simple expression can be applied in fast canopy reflectance models when it is known that the LIDF is approximately spherical.

In section 7.2 it was shown that in the so-called hot spot the bidirectional scattering coefficient w becomes independent of the leaf inclination angle if $\theta_o = \theta_s = 54.7^\circ$. From Fig. 7.9 it appears that small sensitivity to the LIDF can also be reached at other constellations, for instance at $\theta_o = 70^\circ$ with $\theta_s = 35^\circ$ and $\psi = 0^\circ$. Since for the hot spot case the function $T_1(\theta_i)$ was found responsible for the zero sensitivity, one might suspect that the same function would be responsible for the minimum sensitivity to the LIDF outside the hot spot. In that case the function $T_1(\theta_i)$ is given by

$$T_1(\theta_i) = 2\cos^2\theta_i \cos\theta_s \cos\theta_o + \sin^2\theta_i \sin\theta_s \sin\theta_o \cos\psi \quad .$$



This function becomes independent of θ_i if $\tan \theta_s \tan \theta_o \cos \psi = 2$.

For $\psi = 0^\circ$ and $\theta_s = 35^\circ$ this gives $\theta_o = 70.7^\circ$, which is indeed the point in Fig. 7.9 where the sensitivity to the LIDF approximately reaches a minimum. However, the minimum sensitivity found in Fig. 7.9 at $\theta_o = 81^\circ$; $\psi = 180^\circ$ cannot be explained by this. Here the fact that distributions no. 1 (planophile) and no. 9 (erectophile) intersect, probably forms a more significant indication.

Anyway, the observation zenith angles found are too large to be of practical interest, and the extinction coefficient at these angles is still considerably sensitive to the LIDF, so that the resulting canopy reflectance would be also. Therefore, for observation of leaf canopies under minimum sensitivity to the LIDF, the best condition is the hot spot case with $\theta_o = \theta_s = 54.7^\circ$, since in this case the bidirectional scattering coefficient is completely insensitive to the LIDF, while the extinction coefficient shows only little sensitivity.

Table 7.1 Discretized leaf inclination distribution functions.

Angle (deg)	Area fraction in % for distribution no.									
	1	2	3	4	5	6	7	8	9	10
5	72.7	11.8	52.4	0.1	11.1	42.6	0.1	10.4	0.0	1.5
15	12.7	17.6	8.6	0.9	11.1	5.2	0.5	8.0	0.2	4.5
25	6.6	34.2	4.7	3.0	11.1	1.7	1.6	5.5	0.6	7.4
35	3.8	22.1	3.4	9.6	11.1	0.4	3.7	3.9	1.2	10.0
45	2.2	8.4	3.2	72.7	11.2	0.0	8.4	3.2	2.2	12.3
55	1.2	3.7	3.9	9.6	11.1	0.4	22.1	3.4	3.8	14.3
65	0.6	1.6	5.5	3.0	11.1	1.7	34.2	4.7	6.6	15.8
75	0.2	0.5	8.0	0.9	11.1	5.2	17.6	8.6	12.7	16.8
81	0.0	0.1	1.9	0.1	2.2	2.0	2.6	2.8	4.1	3.4
83	0.0	0.0	2.0	0.1	2.2	2.7	2.4	3.6	5.2	3.5
85	0.0	0.0	2.1	0.0	2.3	3.8	2.3	4.9	7.0	3.5
87	0.0	0.0	2.1	0.0	2.2	6.1	2.3	7.6	10.6	3.5
89	0.0	0.0	2.2	0.0	2.2	28.2	2.2	33.4	45.8	3.5



7.4 Angular profiles of bidirectional canopy reflectance

The material of section 7.2 can be (and has been) used for the construction of a very simple vegetation canopy reflectance model, namely one consisting of a single layer of leaves, with given LAI and LIDF, and resting on a Lambertian reflecting soil. The optical properties required are the spectral reflectance and transmittance of the leaves, ρ and τ , and the spectral reflectance of the soil, r_s , at all wavelengths of interest. Further, all assumptions on which the calculation of the SAIL coefficients was based are supposed to be fulfilled (section 7.2). In this case the directional reflectance properties of the canopy layer can be calculated by means of four-stream radiative transfer theory, as presented for instance in Verhoef (1985a). This theory gives the solutions of the system of differential equations (7.1) in the form of four reflectance factors and five transmittance factors, which describe the interaction of incident radiation with the canopy layer as a whole.

The directional reflectance at the top of the canopy is given by

$$r_o = (r_{so}E_{sun} + r_{do}E_{sky}) / (E_{sun} + E_{sky}) \quad ,$$

where E_{sun} = direct solar irradiance

E_{sky} = diffuse sky irradiance

r_{so} = bidirectional reflectance

r_{do} = directional reflectance for diffuse incidence

According to four-stream theory (Verhoef, 1985a), r_{so} and r_{do} are given by

$$r_{so} = \rho_{so} + \frac{(\tau_{ss} + \tau_{sd})r_s(\tau_{oo} + \tau_{do})}{1 - r_s\rho_{dd}} \quad , \text{ and}$$

$$r_{do} = \rho_{do} + \frac{\tau_{dd}r_s(\tau_{oo} + \tau_{do})}{1 - r_s\rho_{dd}} \quad .$$

Here, the double-subscripted ρ and τ quantities are reflectances and transmittances of the isolated canopy layer, and are directly expressed in the SAIL coefficients. The above equations are the result of applying the adding method to the combination canopy-soil and they demonstrate how this technique can be used to obtain more insight in the contributions of soil and canopy to the reflectance of the ensemble.



Angular profiles of the bidirectional reflectance r_{so} in the principal plane have been computed for a number of representative leaf canopies in order to illustrate the most important effects at two wavelengths, namely at 670 nm (red) and 850 nm (near infrared). The optical properties of leaves and soil were taken from Bunnik (1978) and are given by

λ (nm)	ρ (%)	τ (%)	r_s (%)
670	7.5	0.7	17.5
850	52.0	44.0	28.6

These data refer to green wheat leaves and a dry sandy loam soil. Both wavelengths have a good leaf-soil contrast, since in the red the leaves are much darker than the soil, whereas in the near infrared the leaves are much brighter. These wavelengths also represent the ones with maximum and minimum absorption in the leaves (91.8 % in the red and only 4.0 % in the near infrared).

Canopy morphologies have been generated from all combinations of LAI = 0.25; 1.0; 4.0 and LIDF = 2; 4; 7 (see section 7.3). LIDF 4 is the plagiophile distribution, for which most leaf area (72.7 %) has an inclination of 45°. Instead of LIDFs 1 and 9, which are very extreme examples of the planophile and the erectophile type, LIDFs 2 and 7 were chosen as more moderate representatives of these distribution types, having most leaf area inclined at 25° and 65°, respectively. See also Table 7.1 for more details.

The angular profiles as a function of the observation zenith angle θ_o are shown in Figures 7.10 to 7.12 for solar zenith angles θ_s of 30° and 60°. For comparison, (flat) profiles for LAI = 0 (bare soil) have also been drawn. From the profiles simulated by means of the SAIL model one can observe the following features:

- At nadir ($\theta_o = 0^\circ$) all curves have a negative slope, and it becomes stronger with increasing LAI and LIDF number (more inclined leaves). In the neighbourhood of the nadir point the variation with θ_o is almost linear over a range that becomes shorter with the LIDF number. This variation can only be attributed to the bidirectional scattering coefficient w , since this is the only model variable sensitive to the change of observation

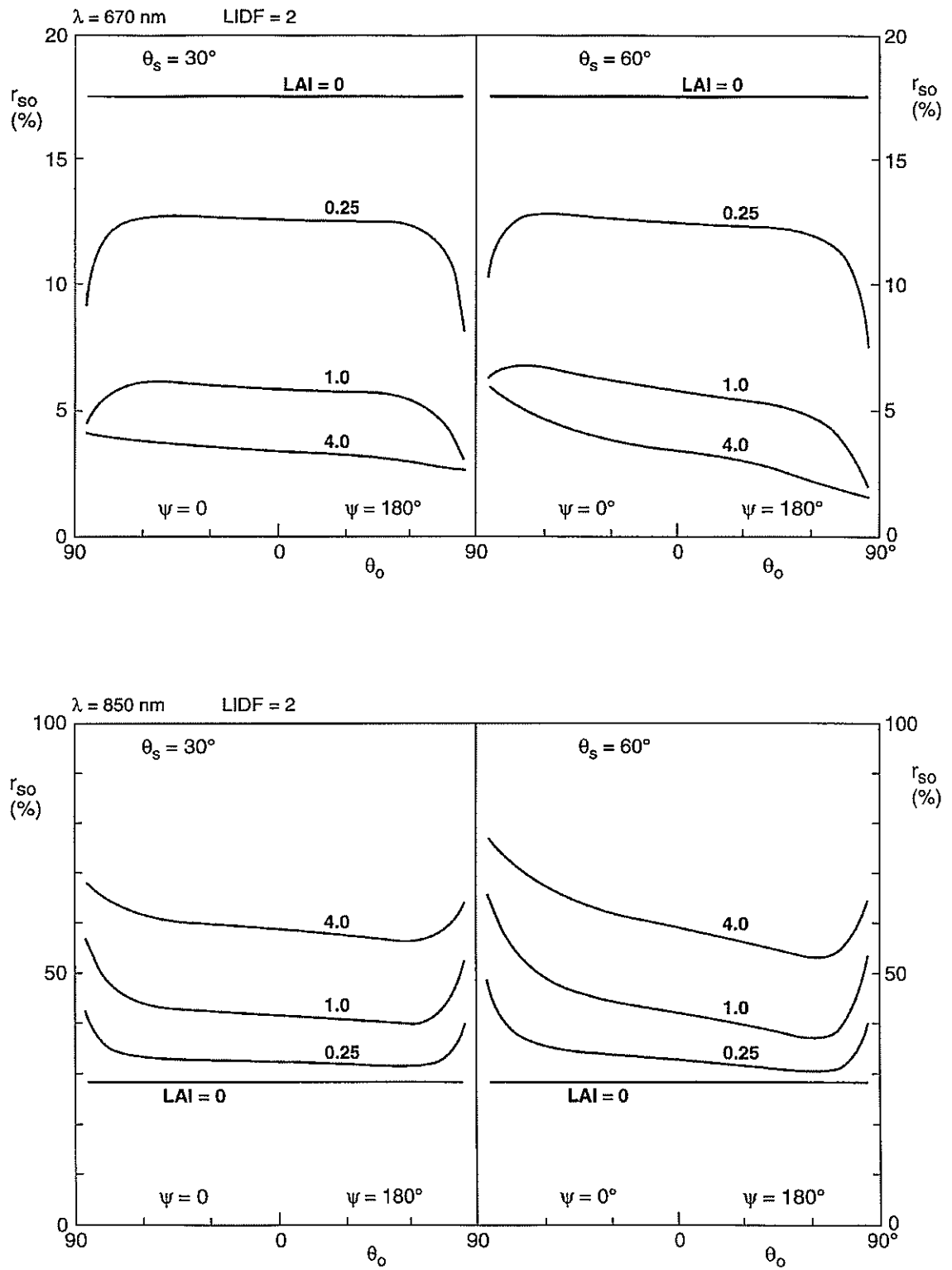


Fig. 7.10 Angular profiles of the bidirectional reflectance for LIDF no. 2

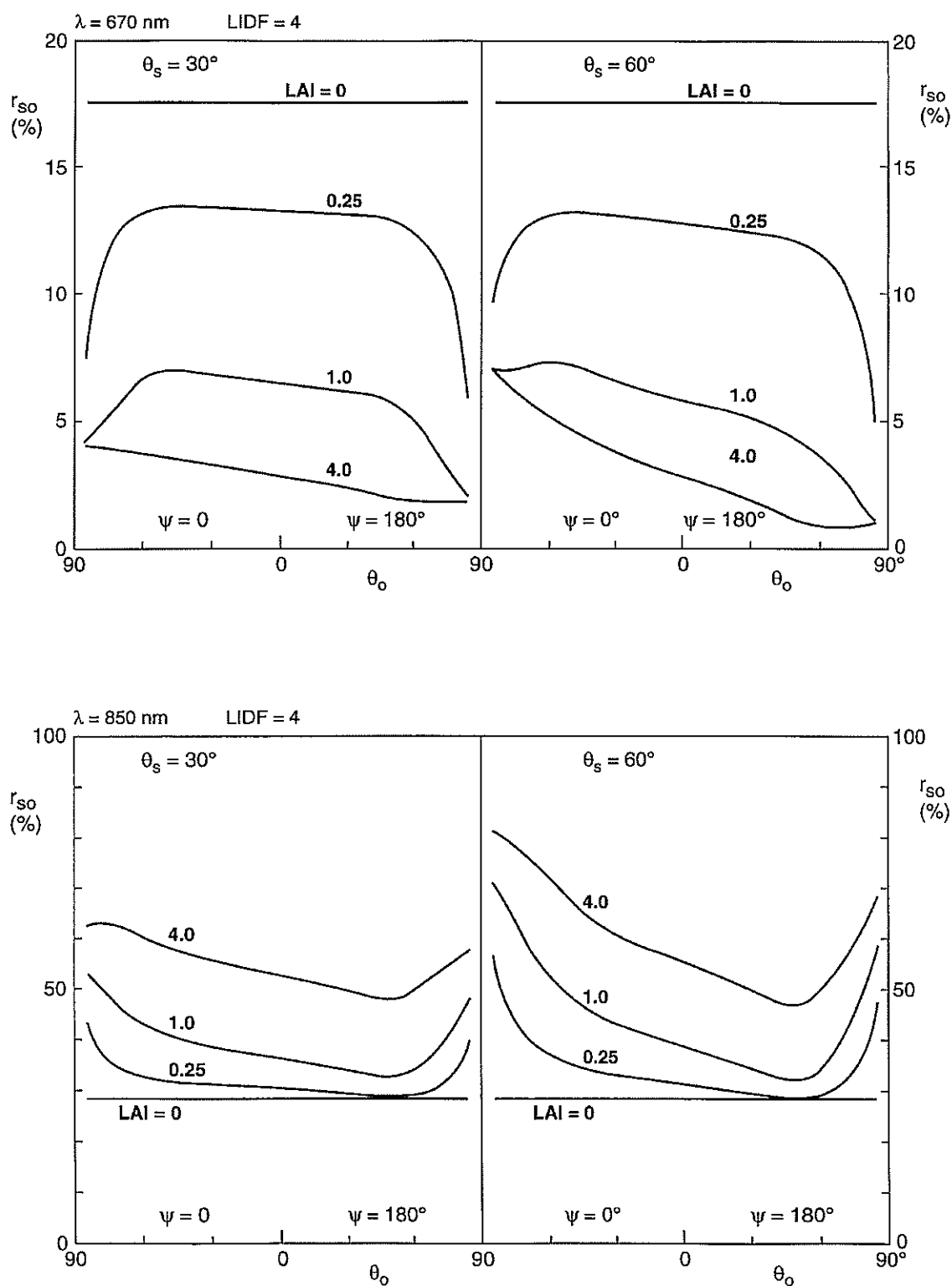


Fig. 7.11 Angular profiles of the bidirectional reflectance for LIDF no. 4

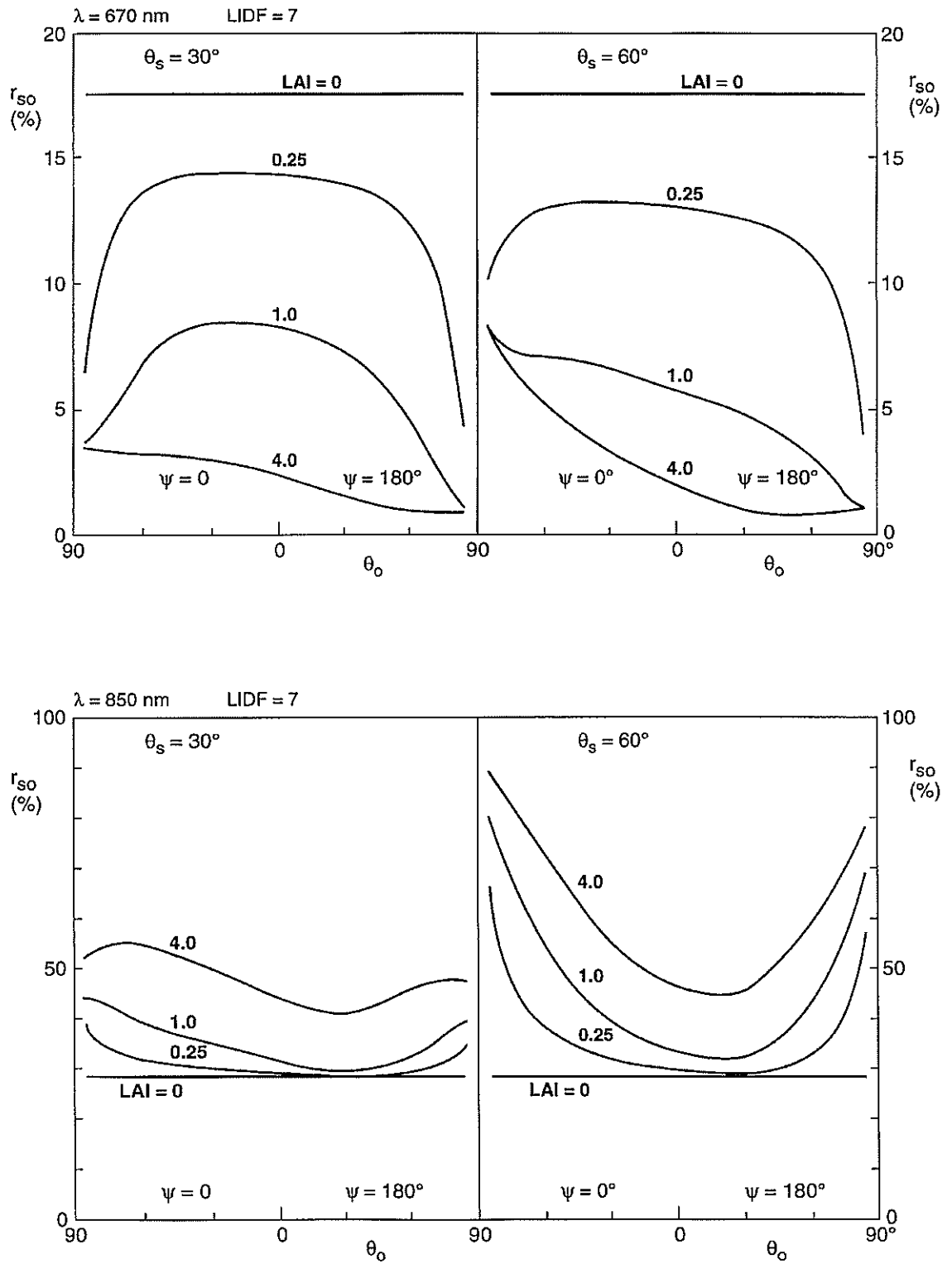


Fig. 7.12 Angular profiles of the bidirectional reflectance for LIDF no. 7



azimuth taking place when one crosses the nadir point.

- For large zenith angles θ_0 , the curves for LAI = 0.25 and LAI = 1.0 bend towards the corresponding curves for LAI = 4.0. This is explained by the increased vegetation cover observed at large viewing angles.
- For low LAI the sensitivity to the LAI is relatively high in the red, at large θ_0 and low LIDF number. At high LAI the sensitivity to changes in the LAI is greatest in the near infrared and low LIDF number, but also in the red a considerable sensitivity is still found, especially for high LIDF number and not too far from nadir.
- The non-Lambertian behaviour of the bidirectional reflectance is most pronounced for large solar zenith angles θ_0 and higher LIDF number. In particular for high LAI and more inclined leaves the influence of the solar zenith angle on the profiles is strong.

The Suits and SAIL models are both based on the assumption that the size of the leaves is negligible, so that the canopy can be treated as a turbid medium. Although the SAIL model forms an improvement over the Suits model in that it gives more realistic directional reflectance profiles for leaf inclination distributions found in reality, the effect of a finite leaf size on the angular profiles is not included, so application of the SAIL model is restricted to vegetation canopies which indeed have very small leaves. A consequence of a finite leaf size is the so-called hot spot effect, i.e. the effect that the amount of shadow observed inside the canopy gradually tends to zero when one approaches the hot spot, the point where the viewing direction coincides with the (opposite) direction of the sunrays.

By means of the SAIL model it is possible to calculate the bidirectional reflectance in the exact hot spot, and it is much higher than in the angular region close to it, so that in the angular profiles presented this would appear as an infinitely sharp peak, which of course is unrealistic. Therefore, the SAIL model has been modified to incorporate this effect in a more realistic manner. This forms the subject of the next chapter.



8 THE HOT SPOT EFFECT IN VEGETATION CANOPIES

8.1 Introduction

The so-called hot spot effect (sometimes called "Heiligenschein" or opposition effect) has been described by various authors, e.g. Minnaert (1940), Suits (1972), Bunnik (1978), Kuusk (1985) and Gerstl (1988), in qualitative as well as quantitative terms. This effect occurs when light enters a 3-D object, is reflected somewhere inside, and subsequently escapes from the object in a direction opposite to the incidence direction. Depending on the internal structure of the object, this escape of the light in opposite direction is much more likely than the escape in other directions, because in the former case the same free path is used twice. The result will be an enhanced brightness of the object in that direction relative to other directions. Vegetation canopies, forests, rough soil surfaces and even cities are examples of objects showing a peak of the reflectance in the retrodirection when exposed to direct sunlight. In general, the occurrence of internal shadows is a good indicator of the possibility that a hot spot effect can be observed.

Of the above mentioned authors, Kuusk gives a physically based description of the hot spot effect that is not limited to the exact hot spot itself. The theory outlined by Kuusk appeared to form a good basis for implementation of the hot spot effect in the SAIL model. The resulting model is called SAILH.

8.2 Theory of Kuusk

Since the hot spot effect as described above is the result of two direct transmissions and one reflection (scattering) event, this is a so-called single scattering phenomenon. In the SAIL model the single scattering contribution to the bidirectional reflectance can be found from the set of differential equations (7.1) of the previous chapter, with all coefficients except w , k and K set to zero. This gives

$$dE_s/dx = k E_s, \quad (8.1.a)$$

$$dE_o/dx = w E_s - K E_o, \quad (8.1.b)$$



in which x is the metrical height (with $x = 0$ at the top) and the coefficients in this case include the leaf area density L' .

These equations describe single scattering in any turbid medium in which the scattering elements are of negligible size. In that case the fluxes E_s and E_o are constant in the horizontal plane. However, when the size of the elements becomes larger, these fluxes will fluctuate in the horizontal plane. For the solar flux E_s these fluctuations are simply formed by the presence or absence of sunlight. So, if Eqs. (8.1) are still valid in this case, then the fluxes can only refer to average values over an area much larger than the size of the scattering elements.

In the solution of Eq. (8.1.a), which is given by $E_s(x) = E_s(0)e^{kx}$, the term e^{kx} can be interpreted as a probability P_s of the presence of sunlight in the canopy. This interpretation remains correct in the case of a finite leaf size, provided the spatial distribution of the leaves is uniform (at random).

For the solution of Eq. (8.1.b) one can first differentiate $E_o e^{Kx}$, which yields

$$\frac{d}{dx} E_o e^{Kx} = K E_o e^{Kx} + e^{Kx} \frac{dE_o}{dx}.$$

Substitution of Eq. (8.1.b) then gives

$$\frac{d}{dx} E_o e^{Kx} = w E_s e^{Kx} = w E_s(0) e^{(K+k)x}. \quad (8.2)$$

Here the term $e^{(K+k)x}$ is interpreted as the joint probability of the presence of sunlight and free line of sight from outside the canopy. This probability is called P_{so} , and it is the product of $P_s = e^{kx}$ and $P_o = e^{Kx}$, where the latter is the probability of free line of sight.

The solution of Eq. (8.2) is given by

$$E_o(x) e^{Kx} = \int_{-h}^x w E_s(0) e^{(K+k)x} dx + r_s E_o(0) e^{-(K+k)h},$$

where r_s is the reflectance of the soil and h is the canopy height. In terms of the probabilities introduced above, this solution is rewritten as



$$E_o(x)P_o(x) = \int_{-h}^x w E_s(0) P_{so}(x) dx + r_s E_s(0) P_{so}(-h) \quad .$$

For the single scattering contribution to the bidirectional reflectance, called r_{so}^s , one finds

$$r_{so}^s = w \int_{-h}^0 P_{so}(x) dx + r_s P_{so}(-h) \quad . \quad (8.3)$$

The hot spot effect is now easily explained by the fact that the probabilities P_s and P_o are no longer independent of one another under the hot spot condition, so that P_{so} is much greater than the product $P_s P_o$. In the exact hot spot, $P_{so} = P_s = P_o$, whereas only far outside the hot spot $P_{so} = P_s P_o$.

In both these extreme cases Eq. (8.3) can easily be solved analytically, cf. Suits (1972), Bunnik (1978).

The theory of Kuusk (1985) also provides a solution for the intermediate case. According to this theory, the probabilities P_s and P_o are the result of a large number of events, s and o , which occur with probabilities p_s and p_o , respectively. These events are the penetration of light through thin layers Δx in the direction from the sun and in the direction of observation. The probabilities of these elementary events are given by

$$p_s = 1 - k \Delta x \quad \text{and} \quad p_o = 1 - K \Delta x \quad .$$

Now consider a sunray which penetrates into the canopy down to a level x , is reflected by a leaf, and escapes through the canopy in the direction of viewing. Figure 8.1 illustrates this situation.

Let the depth from the top of the canopy be z . Then the horizontal distance between the two rays at the top is given by

$$d(0) = z \sqrt{\tan^2 \theta_s + \tan^2 \theta_o - 2 \tan \theta_s \tan \theta_o \cos \psi} = z \alpha \quad .$$

At level $x = -z$ the rays intersect, so here the horizontal distance between them is zero. In general,

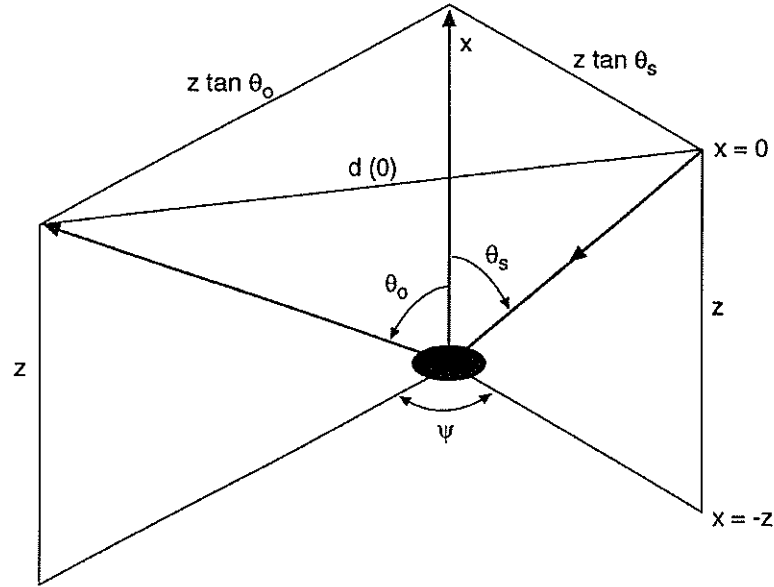


Fig. 8.1 Path of a reflected sunray in a leaf canopy

$$d(x) = (x+z)\alpha .$$

When the distance $d(x)$ becomes small, the events s and o at level x will be correlated if the leaf size is finite. It seems reasonable to assume that this correlation varies exponentially as

$$\rho = e^{-d(x)/\ell} ,$$

where ℓ is the so-called horizontal correlation length. This correlation length depends on the leaf size, but also on leaf shape and inclination. However, this relationship is very complex, so it is simply assumed that ℓ is a given parameter. Computational techniques like ray-tracing might be used in order to investigate how ℓ depends on the canopy architecture. Returning now to the elementary events s and o , it can be assumed that these events occur according to a Poisson-distribution. This can also be assumed for the combined event (s and o) with the associate probability p_{so} . Statistically, the correlation coefficient for the events s and o follows from



$$\rho = \frac{\text{covar}(s \text{ and } o)}{\sqrt{\text{var}(s) \text{var}(o)}} ,$$

$$\begin{aligned} \text{where } \text{covar}(s \text{ and } o) &= p_{so} - p_s p_o , \\ \text{var}(s) &= p_s(1-p_s) , \text{ and} \\ \text{var}(o) &= p_o(1-p_o) , \text{ so} \end{aligned}$$

$$\rho = \frac{p_{so} - p_s p_o}{\sqrt{p_s p_o (1-p_s) (1-p_o)}} , \text{ or} \quad (8.4)$$

$$p_{so} = p_s p_o + \rho \sqrt{p_s p_o (1-p_s) (1-p_o)} .$$

For small Δx , such that $k \Delta x \ll 1$ and $K \Delta x \ll 1$, this can be approximated by

$$p_{so} = 1 - (K+k) \Delta x + \rho \sqrt{k \Delta x K \Delta x} = 1 - (K+k - \rho \sqrt{Kk}) \Delta x .$$

Substitution of ρ and $d(x)$ gives

$$p_{so}(x) = 1 - [K+k - \sqrt{Kk} e^{-(x+z)\alpha/\ell}] \Delta x .$$

The total joint probability $P_{so}(x)$ required by Eq. (8.3) is found from the product of all probabilities $p_{so}(x)$ over all levels from level $-z$ to the top of the canopy. This product can be written as an integral if $p_{so}(x)$ is expressed as an exponential:

$$\begin{aligned} p_{so}(x) &= \exp [- \{ K+k - \sqrt{Kk} e^{-(x+z)\alpha/\ell} \} \Delta x] , \text{ and} \\ P_{so}(z) &= \exp [\int_{-z}^0 - \{ K+k - \sqrt{Kk} e^{-(x+z)\alpha/\ell} \} dx] \\ &= \exp [-(K+k)z + \sqrt{Kk} \frac{\ell}{\alpha} (1 - e^{-z\alpha/\ell})] , \end{aligned}$$

where the joint probability is expressed in the depth z because the integration was over x from $-z$ to 0 . It can also be expressed in x by



$$P_{so}(x) = \exp \left[(K+k)x + \sqrt{Kk} \frac{\ell}{\alpha} (1 - e^{x\alpha/\ell}) \right] \quad (8.5)$$

For small ℓ/α this expression approaches $\exp[(K+k)x]$, which is the familiar expression for the joint probability far away from the hot spot or for very small leaves. For large ℓ/α one can write

$$1 - e^{x\alpha/\ell} \approx 1 - (1 + x\alpha/\ell) = -x\alpha/\ell, \text{ so that one obtains}$$

$$P_{so}(x) \approx \exp \left[(K+k)x - \sqrt{Kk} x \right]$$

However, this can only occur if α is very small, which means that the directions of sun and observer are almost coincident, so that K and k are practically equal, and one obtains $P_{so}(x) = \exp(Kx) = \exp(kx)$.

It can thus be concluded that Eq. (8.5) gives correct approximations of the joint probability in both extreme cases.

So far this discussion of Kuusk's theory. It should be mentioned that the terminology and notation have been adapted to the style used elsewhere in this manuscript. Also, Kuusk considers more possibilities for the form of the correlation function than just the exponential one. However, he concludes that this does not affect the results very much.

A difficulty with Eq. (8.5) arises if one considers a small value of x , say $x = -\delta$. In that case the result is an expression identical to the one just found for the near hot spot, but now this result is obtained regardless of the ratio ℓ/α , so K and k are not necessarily equal (are almost equal). The result becomes

$$P_{so}(-\delta) = \exp \left[-(K+k - \sqrt{Kk})\delta \right]$$

Now, suppose K is considerably greater than k . In that case $K > (Kk)^{1/2} > k$, so that $k - (Kk)^{1/2}$ is negative, which leads to the conclusion that $P_{so}(-\delta)$ is greater than $P_o(-\delta)$. However, this is statistically impossible, since the joint probability cannot be greater than any of the individual probabilities P_s and P_o . The maximum of P_{so} that can be attained is equal to the minimum of P_s and P_o . In his paper, Kuusk did not address this problem, and one might also argue that if δ is small, then P_{so} is close to one anyway, so that the effect on the final result would only be minor. Nevertheless, a slight



modification, namely the replacement of the square root term $(Kk)^{1/2}$ by the minimum of K and k , would be sufficient to cure the problem. This can also be justified on statistical grounds by reconsidering Eq. (8.4) in the case of maximum correlation when p_s and p_o are different.

Since the maximum of p_{so} equals the minimum of p_s and p_o , the maximum correlation is given by

$$\rho_{\max} = \frac{\min(p_s, p_o) - p_s p_o}{\sqrt{p_s p_o (1 - p_s)(1 - p_o)}} .$$

For small Δx this can be approximated by

$$\rho_{\max} = \min [\sqrt{K/k}, \sqrt{k/K}] .$$

The dependence of the correlation coefficient on the horizontal distance $d(x)$ should now be expressed by

$$\rho = \rho_{\max} e^{-d(x)/\ell} .$$

Substitution of this in the equation for $p_{so}(x)$ gives

$$p_{so}(x) = 1 - [K + k - \min(K, k) e^{-(x+z)\alpha/\ell}] \Delta x .$$

For Eq. (8.5) this indeed leads to the replacement of $(Kk)^{1/2}$ by $\min(K, k)$.

The terms $(Kk)^{1/2}$ and $\min(K, k)$ are both examples of estimates of a more general function Λ , which expresses the average maximum overlap between the leaf projections on a horizontal plane for the directions of the sun and of viewing. By maximum overlap is meant the overlap of the projections when there is no shift due to parallax, i.e. the leaves are projected on the horizontal plane through the centre of the leaves. For horizontal leaves this gives $\Lambda = K = k$, so both estimates produce the same correct answer. However, for inclined leaves the maximum overlap will be less than that.

For instance, inclined circular leaves produce projections in the form of ellipses, and when the directions of the sun and of viewing are different, the overlap of the ellipses is less than either of the elliptical projections associated with the sun and viewing, simply because the ellipses do not completely match. Therefore it can be concluded that even the function



$\min(K, k)$ in general will overestimate the overlap. The consequence is that the hot spot effect outside the actual hot spot is also overestimated, in particular for Kuusk's overlap function. However, numerical experiments (cf. section 8.5) have shown that more accurate calculation of the overlap function Λ based on circular leaves makes the entire model one hundred times slower, so for simulations of a somewhat less precision Kuusk's overlap function might still be preferred.

8.3 Incorporation into the SAIL model

As mentioned in the previous chapter (section 7.4), the SAIL model uses dimensionless coefficients of extinction and scattering, with an associated relative optical height as the vertical co-ordinate. In this way the canopy height h becomes a redundant parameter. In the previous section, however, a metrical height was used for the description of the hot spot effect, so there the coefficients w , k and K have a dimension of m^{-1} . Introduction of the relative optical height in the most important equations of section 8.2 proceeds as follows:

Eq. (8.5) is written as a function of x/h by

$$P_{so}(x/h) = \exp \left[(K+k)hx/h + \sqrt{Khkh} \frac{\ell}{h\alpha} \left(1 - e^{\frac{x}{h} \frac{h\alpha}{\ell}} \right) \right] .$$

Similarly, Eq. (8.3) is rewritten as

$$r_{so}^s = wh \int_{-1}^0 P_{so}(x/h) d(x/h) + r_s P_{so}(-1) .$$

By means of the transformations $x/h \rightarrow x$, $wh \rightarrow w$, $kh \rightarrow k$ and $Kh \rightarrow K$, both equations are rewritten as

$$P_{so}(x) = \exp \left[(K+k)x + \sqrt{Kk} \frac{\ell}{h\alpha} \left(1 - e^{\frac{x}{h} \frac{h\alpha}{\ell}} \right) \right] , \text{ and} \quad (8.6)$$



$$r_{so}^s = w \int_{-1}^0 P_{so}(x) dx + r_s P_{so}(-1) \quad , \quad (8.7)$$

where w , k , K and x are now dimensionless.

A new dimensionless parameter obtained by this transformation is the ratio ℓ/h , which is called the hot spot size parameter. This is the only additional parameter necessary to describe the hot spot effect for a single layer canopy. As the integral of $P_{so}(x)$ over the canopy layer has no analytical solution, it is determined by means of a numerical algorithm. For this purpose, the range of x from -1 to 0 is divided in 20 unequal intervals Δx in such a way that equal steps in the function $e^{x\gamma}$, where $\gamma = h\alpha/\ell$, are obtained. Subsequently, $P_{so}(x)$ is assumed to vary purely exponentially in each interval, so that the analytical integral over each interval can easily be determined. The total integral is approximated by a summation as

$$\int_{-1}^0 P_{so}(x) dx \approx \sum_{i=1}^{20} \frac{\Delta e^{y_i}}{\Delta y_i} \Delta x_i \quad ,$$

where $y_i = (K+k)x_i + \sqrt{Kk} \frac{1-e^{x_i\gamma}}{\gamma}$.

In the terminology of four-stream theory, the single scattering contribution to the bidirectional reflectance can be written as

$$r_{so}^s = \rho_{so}^s + \tau_{ssoo} r_s \quad ,$$

$$\text{where } \rho_{so}^s = w \int_{-1}^0 P_{so}(x) dx \quad , \text{ and } \tau_{ssoo} = P_{so}(-1) \quad .$$

Here τ_{ssoo} is the joint two-way direct transmittance (sun-soil-observer) through the entire canopy layer. Sometimes this quantity is called the bidirectional gap probability.

Apart from the single scattering contribution there is still also a contribution due to multiple scattering from the canopy. This contribution is called ρ_{so}^d and it is computed as before, but with exclusion of w from the equations. The sum of both contributions gives the bidirectional reflectance from the

canopy as

$$\rho_{so} = \rho_{so}^s + \rho_{so}^d \quad .$$

For the bidirectional reflectance of a single layer canopy on a soil background with reflectance r_s one finds

$$r_{so} = \rho_{so} + \left[\tau_{ssoo} + \frac{(\tau_{sd} + \tau_{ss} r_s \rho_{dd}) \tau_{oo} + (\tau_{ss} + \tau_{sd}) \tau_{do}}{1 - r_s \rho_{dd}} \right] r_s \quad . \quad (8.8)$$

From a software implementation point of view, the incorporation of the hot spot effect according to Kuusk means only one extra input parameter, namely ℓ/h , and one extra output parameter, namely the joint two-way transmittance τ_{ssoo} .

Eq. (8.8) can be applied for calculation of the bidirectional reflectance of the combination vegetation-soil. This equation is still rather simple because the adding algorithm could be applied. However, this is only the case for a single canopy layer. If there are more layers in which the hot spot effect is modelled, then the interaction becomes more complex, and a simple adding algorithm can no longer be used. In that case one has to apply the boundary equation method for the coupling of the layers.

8.4 Crop growth considerations

From the previous sections one might conclude that incorporation of the hot spot effect is primarily intended as a correction of the canopy reflectance with regard to bidirectional effects. However, the introduction of the new hot spot size parameter ℓ/h also forces one to consider the development of this parameter during the growth of a crop. Can it be considered constant for certain crop types or does it change during the growing season?

In the first place, crop growth can be characterized by an increasing leaf area index L . More leaf area means more interception of visible radiation which is used for photosynthesis. For making more leaf area a plant can follow two basically different strategies: 1) increasing the number of leaves per unit ground area by making new leaves or 2) letting the existing leaves grow in size. The first strategy is roughly followed by for instance wheat,



corn and potatoes. The second strategy is followed by for instance sugar beet. However, in reality most crop types will probably follow a mixture of both strategies, and it is not unlikely that the partitioning over both can change with the growth phase.

Strategy one is usually accompanied by an increasing height of the crop, because new leaves shoot - at regular intervals - from the stem which grows in vertical direction, such as in wheat. For such a type of crop it may be assumed that the canopy height grows proportionally to the leaf area index. If one assumes that the linear leaf size is more or less constant, then one obtains the relationship $\ell/h \propto 1/L$.

For strategy two it may be assumed that the leaves grow in area, but have a constant shape. In this case the height of the canopy is mainly determined by the linear size of the leaves (such as for sugar beet crops), which means that the ratio ℓ/h is constant. In the next section some results of simulations of the hot spot effect for both growth strategies are presented.

8.5 Examples of simulations

For the visible part of the spectrum, a solar zenith angle of 45° and observation in the principal plane Fig. 8.2 shows the bidirectional reflectance profiles for a series of leaf area indices for growth strategy one. The leaf angle distribution is spherical and the size parameter $\ell/h = 0.1/L$. The hot spot effect is clearly visible by the sharp peaks in the bidirectional reflectance profiles. In this case the hot spot size parameter is large for small L , which leads to a broad peak. For higher L the peak becomes sharper because of the decreasing size parameter. In this part of the spectrum the reflectance is very sensitive to the leaf area index L when it is small. However, saturation occurs above $L = 4$, as expressed in Fig. 8.2 by the small difference with the curve for $L = 8$. The greatest difference between both curves is located at the hot spot, which is caused by the fact that here the combined (two-way) extinction coefficient reaches a minimum.

For otherwise the same conditions as in Fig. 8.2, Fig. 8.3 shows the bidirectional profiles for a fixed hot spot size parameter equal to 0.1. In this case the hot spot peak is sharper than in Fig. 8.2 for $L < 1$ and less sharp for $L > 1$. As a function of the leaf area index a remarkable behaviour is displayed. At most viewing angles a minimum reflectance is reached at $L = 4$, and then it starts to increase. Saturation does not occur, except in the

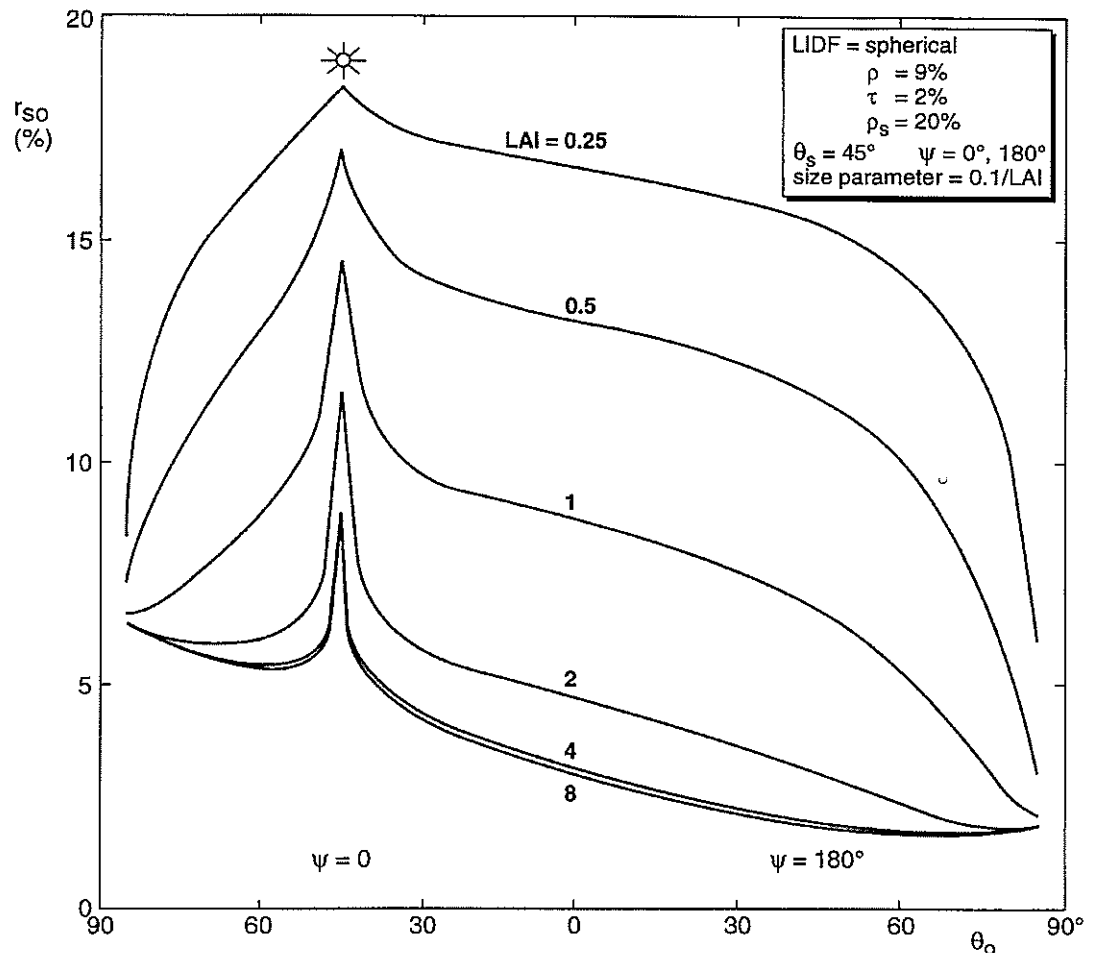


Fig. 8.2 Hot spot effects in the visible for growth strategy 1

exact hot spot, where the behaviour is the same as in Fig. 8.2. Even the step from $L = 16$ to $L = 32$ gives a significant increase of the reflectance outside the actual hot spot.

A physical explanation for this model-predicted phenomenon can be given by considering a leaf layer of constant thickness and leaf size. When in that case the leaf area index is increased, the leaf area density also increases. One can imagine that for a high leaf area density there will be a limit to the depth to which light can penetrate into the canopy. Parts of the canopy located below this level will not contribute to the reflectance and it would make no difference if this part of the canopy would be removed. However, in that case the penetration depth should be taken as the measure of canopy height in the definition of the hot spot size parameter, not the original canopy thickness. Since for high leaf area densities the penetration depth is only a fraction of

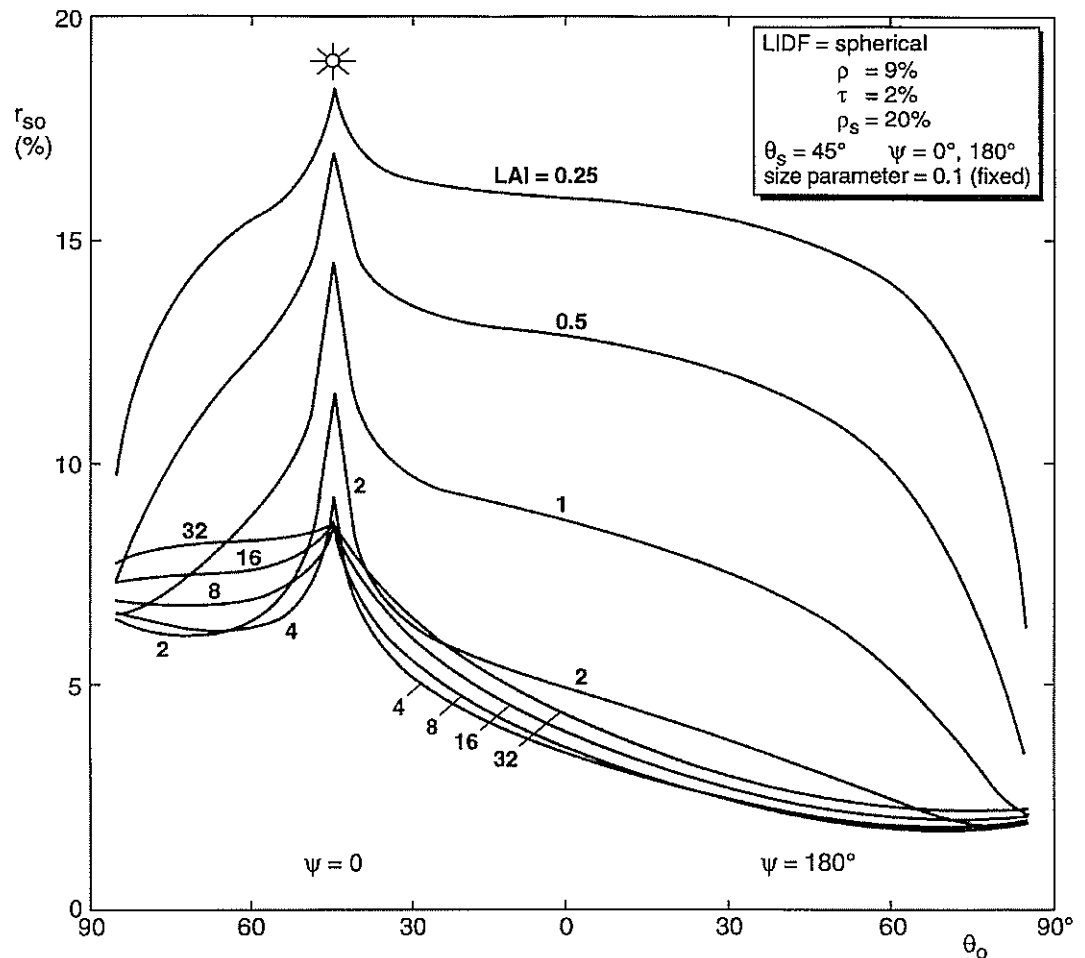


Fig. 8.3 Same as Fig 8.2 but growth strategy 2

the total canopy thickness, the hot spot size parameter becomes much greater than the one based on the canopy height, which effectively leads to an associate broadening of the hot spot peak. As a matter of fact, this is exactly what is observed in Fig. 8.3: a broadening of the hot spot zone for high leaf area indices, leading to increasing reflectances outside the actual hot spot, but not in the hot spot itself.

It is also interesting to investigate the above behaviour for the near infrared part of the spectrum, since this region is mostly used (in combination with the red) for estimation of the leaf area index. Fig. 8.4 shows the results for the near infrared under conditions otherwise equal to those of Fig. 8.3. In this case the hot spot effect is less pronounced because in this wavelength region multiple scattering also plays a significant role in the reflectance. Here the reflectance is a monotonously increasing function of the leaf area

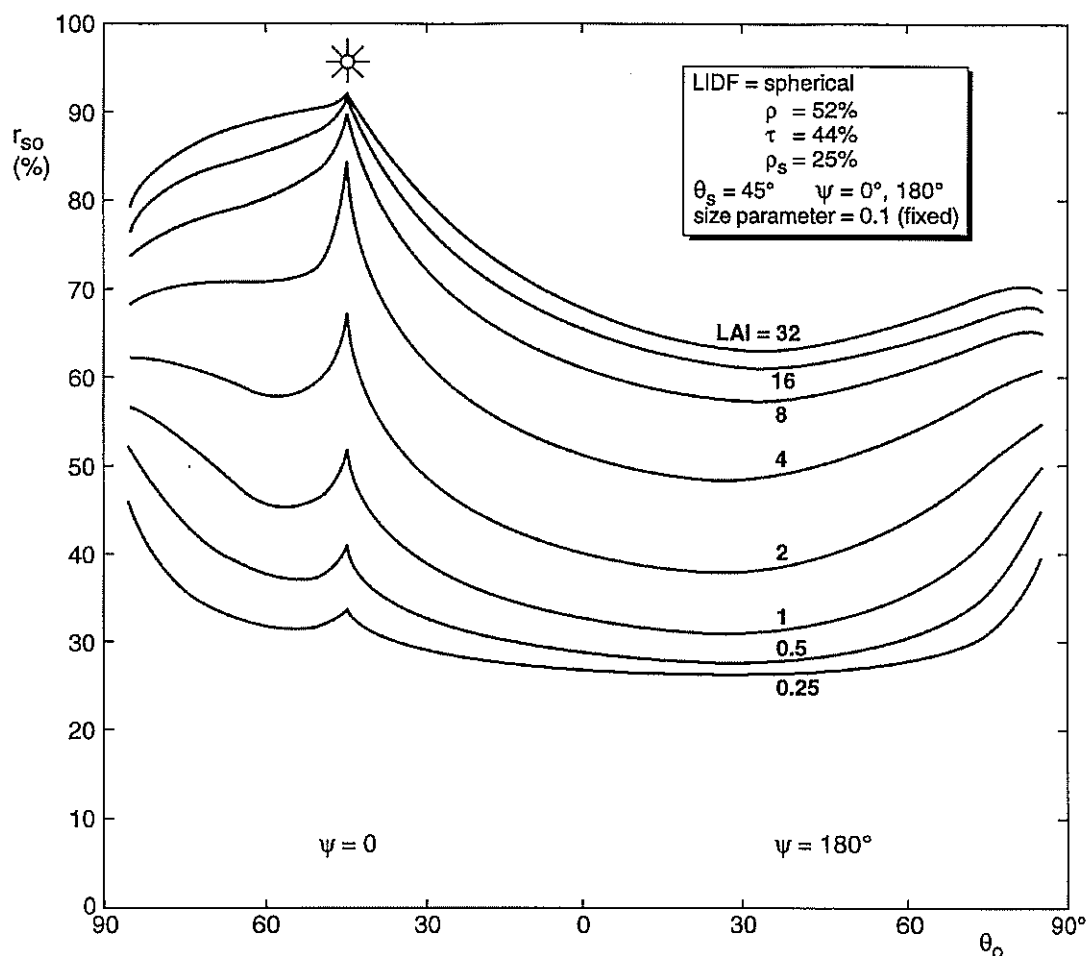


Fig. 8.4 Hot spot effects in the near infrared for growth strategy 2

index and again no saturation occurs because of the above mentioned phenomenon. In the SAIL model the leaf size is infinitesimally small and in that model saturation does occur at a leaf area index of about 8, thus leading to statements on the limited possibilities for estimating the leaf area index from remotely sensed data. However, Fig. 8.4 indicates that for crops following growth strategy two the reflectance remains sensitive to the leaf area index even when it has reached levels greatly exceeding the plausible biological range.

In section 8.2 a theoretical limitation of Kuusk's overlap function was mentioned, namely that it would cause some overestimation of the hot spot effect. In Figures 8.5 and 8.6 it is demonstrated how serious this is in the red part of the spectrum for four different canopy architectures. For some

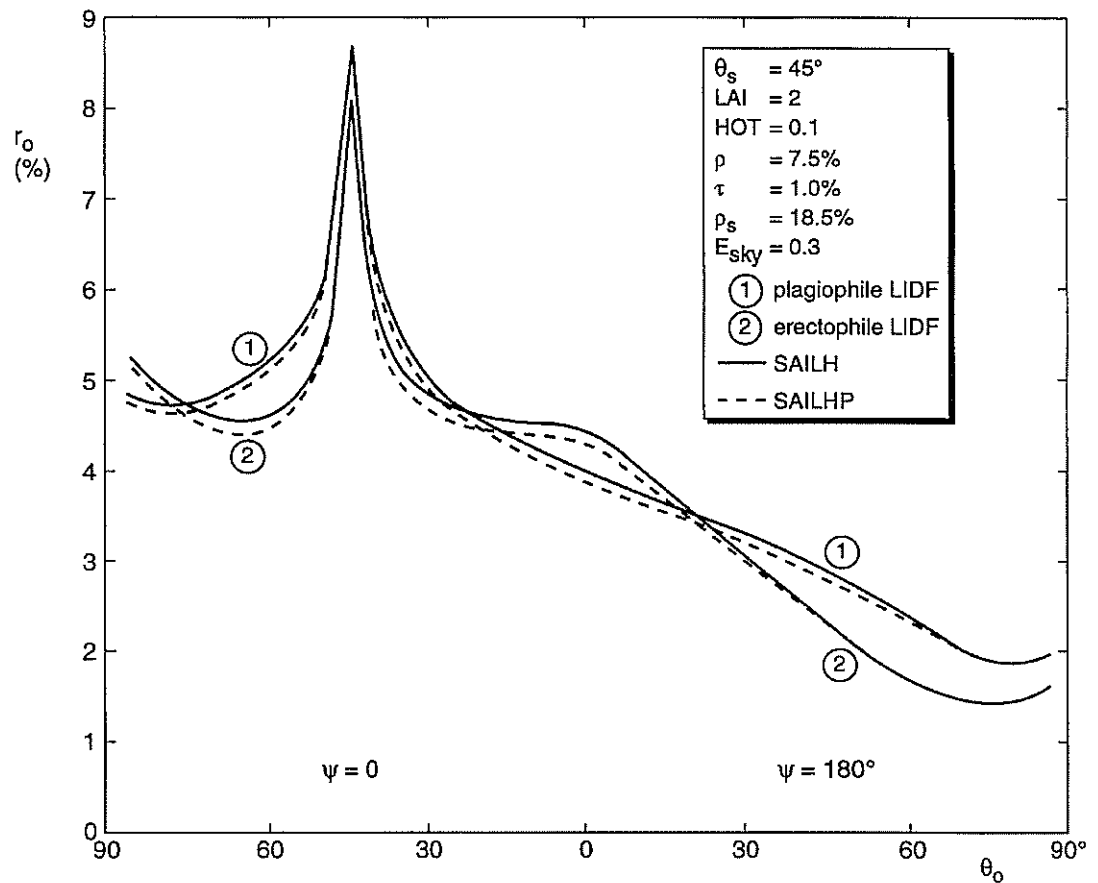


Fig. 8.5 Effects of precise calculation of the overlap function for a size parameter of 0.1 and LAI = 2

more realism, in this case the directional reflectance factor r_o is shown in the principal plane for $E_{sky} = 0.3$ (30 % of incident radiation is diffuse) and a solar zenith angle of 45° . In Fig. 8.5 the leaf area index is 2 and the hot spot size parameter is 0.1. Two leaf angle distributions are shown, namely plagiophile (LIDF no. 4 of chapter 7) and erectophile (LIDF no. 9). The solid lines give the results for the SAILH model with Kuusk's overlap function and the dashed lines those obtained for the experimental model SAILHP ("high precision"), specially developed for this occasion. In the latter model the leaf projection overlap function is calculated numerically under the assumption of inclined circular leaves, at the price of a hundred-fold increased execution time. The results indicate that Kuusk's overlap function indeed overestimates the hot spot effect, but the differences are not large. The greatest differences occur for the erectophile leaf angle distribution and in that case are about 0.2 % (absolute reflectance difference)

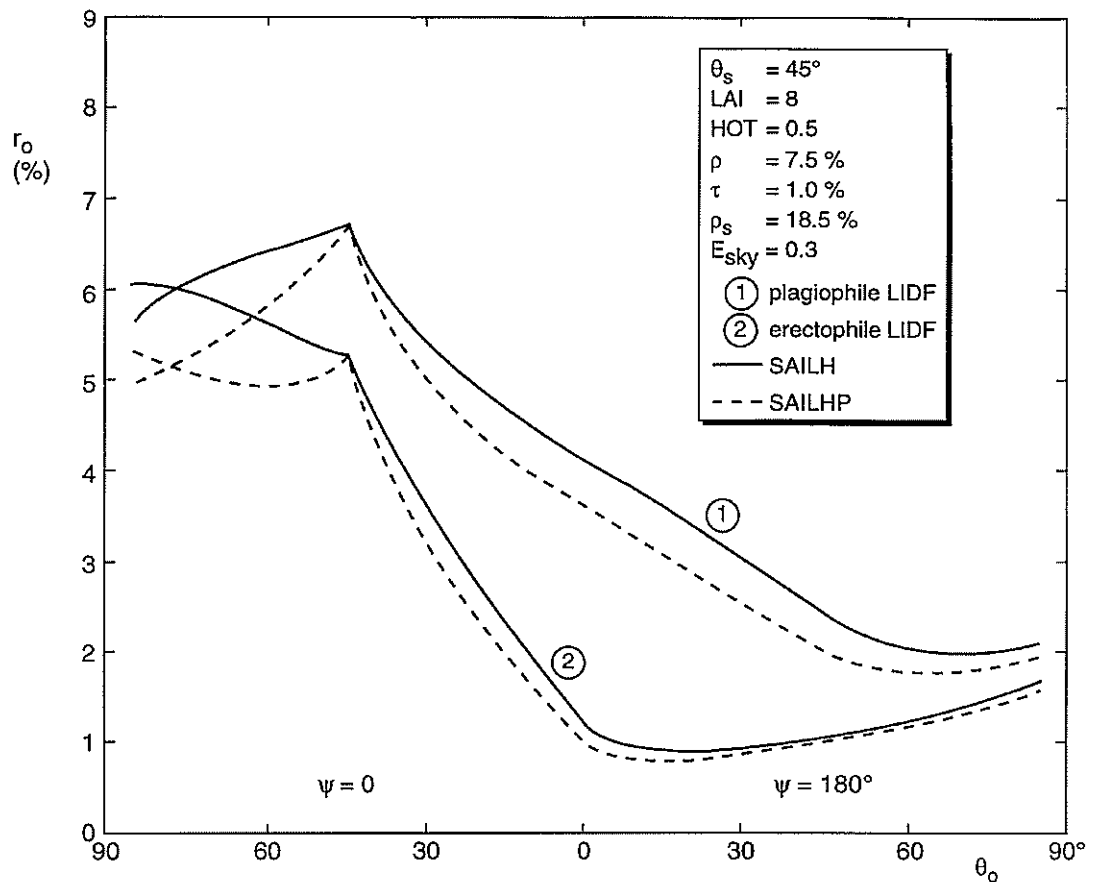


Fig. 8.6 Same as Fig 8.5 but for size parameter of 0.5 and LAI = 8

at maximum. For vertical viewing the difference is 0.1 % for both leaf angle distributions.

For a more extreme case, namely a leaf area index of 8 and a size parameter of 0.5, but otherwise equal conditions, the results are presented in Fig. 8.6. Here the maximum absolute difference is almost 1 % and relative differences of 10 to 20 % are found. However, it remains questionable whether this combination of canopy parameters is biologically plausible, as this case corresponds to a very high leaf area density.

From the above results it can be concluded that under usual circumstances the use of Kuusk's overlap function might still be preferred, since the errors are not great in comparison to other possible sources of error, and the alternative is a model which runs one hundred times slower.



9 FOUR-STREAM ATMOSPHERIC CORRECTION MODEL

9.1 Introduction

In this chapter the effect of the atmosphere on optical remotely sensed images is described by means of four-stream radiative transfer theory. Although in chapter 6 it was shown that an $(N+2)$ -stream version is more accurate, the uncertainty about the true values of many of the input parameters of an atmospheric model usually is so great that application of a more accurate model would hardly be justified.

For the atmosphere, approximate extinction and scattering coefficients are derived based on aerosol scattering, Rayleigh scattering and absorption by water vapour and ozone gas. These are subsequently applied in order to obtain the reflection and transmission properties of the atmospheric layer. By means of the adding method the bidirectional reflectance of the combination atmosphere - earth's surface can be found. This quantity is also called planetary reflectance and the signal detected by earth observation satellites in the "optical" window ($0.4 - 2.5 \mu\text{m}$) is directly proportional to it.

Atmospheric correction is the derivation of the reflectance of the earth's surface from the planetary reflectance. This can be useful for radiation budget studies, for a better spectral characterization of objects on the ground and for the monitoring of object parameters independently of changing atmospheric conditions. The conditions under which atmospheric correction can be carried out successfully are limited, however. The main requirements are that the atmosphere is laterally homogeneous, its constituents are known, and that the earth's reflectance can be approximated as being Lambertian.

Since in most cases the concentrations of some atmospheric constituents, such as water vapour and aerosol, are not known, one usually applies techniques to estimate these quantities from the imagery or from meteorological observations. The less variable effects, such as Rayleigh scattering, are described in the literature and can easily be included in atmospheric models.

9.2 Description of the atmospheric effect

The effect of the atmosphere on satellite observations of the earth is illustrated in Fig. 9.1. Here one can identify three contributions to the radiance detected by the satellite: 1) a contribution from the target il-

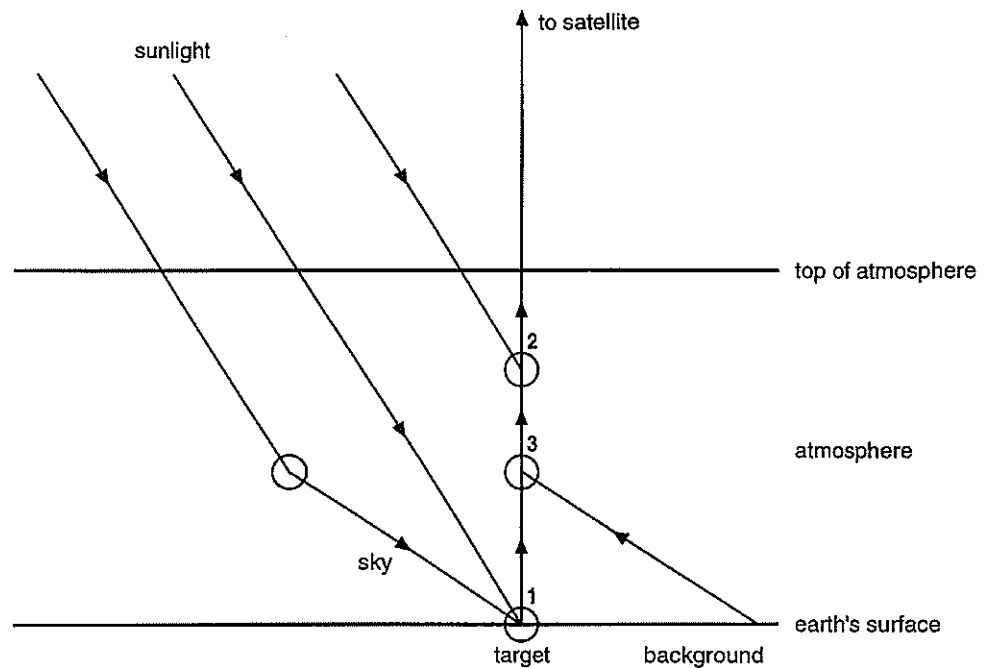


Fig. 9.1 Atmospheric effect on satellite observations of the earth

luminated by direct sunlight and diffuse skylight and 2) a contribution caused by scattered sunlight and 3) a contribution from objects outside the field of view. The latter two contributions together form the so-called path radiance, in which 2) is the atmospheric part and 3) the "background" part. The background contribution is said to be caused by the "adjacency"-effect, expressing that objects in the neighbourhood of the target also contribute to the detected radiance. An equation for the radiance detected at the satellite is given by

$$\pi L_s = \pi L_p + E_{tot} r_t T , \quad (9.1)$$

where L_s = radiance at satellite
 L_p = path radiance
 E_{tot} = total irradiance on the target
 r_t = target reflectance (assumed Lambertian)
 T = target-satellite transmittance

In terms of four-stream theory the atmospheric effect is illustrated in Fig. 9.2, which is an example of a so-called flux interaction diagram.

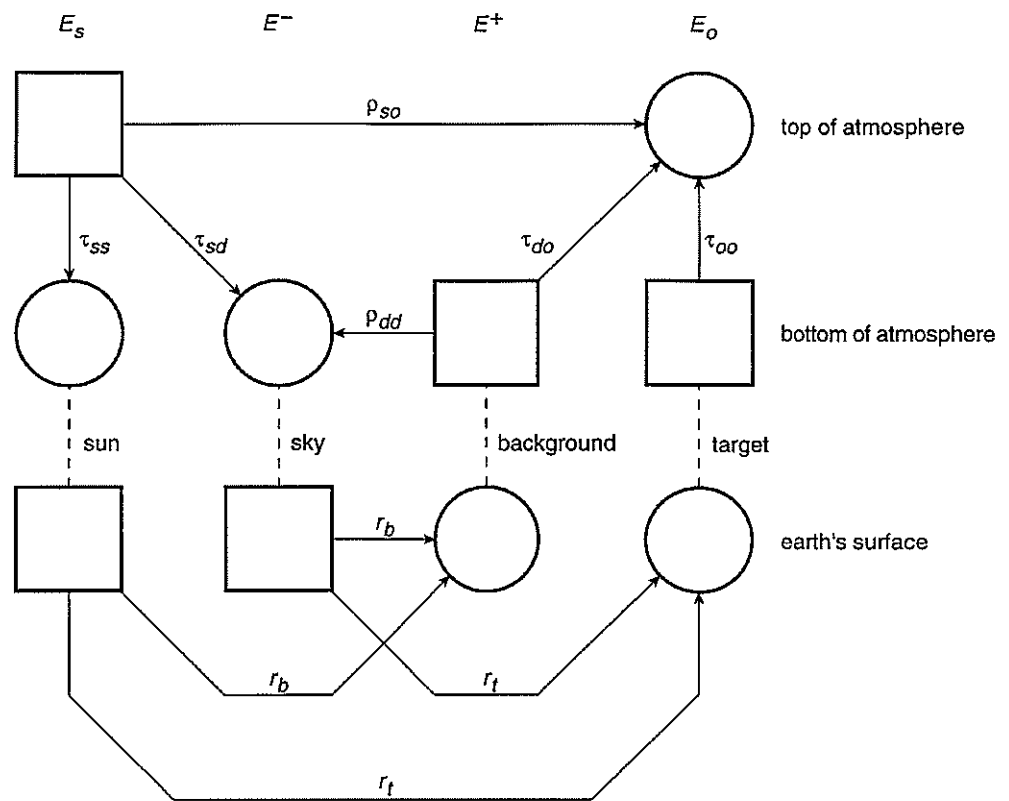


Fig. 9.2 Four-stream flux interaction diagram for the atmospheric effect

In this diagram each incident flux is represented by a square and each exitent flux by a circle. Each arrow indicates the direction of flow and the associated quantity is a reflectance factor (ρ or r) or a transmittance factor (τ). The subscripts for the atmospheric quantities refer to the types of incident and exitent flux, i.e. s for direct solar flux, d for diffuse flux and o for flux in the observer's direction. For the earth's surface the subscripts t and b refer to target and background, respectively. At the interface atmosphere - earth's surface the downward fluxes at the bottom of the atmosphere are exitent from the atmosphere and incident to the earth's surface at the same time, and a similar situation holds for the upward fluxes. In Fig. 9.2 dashed lines are used to express these identities. Next to the dashed lines the different contributions at ground level are mentioned. The transfer equations associated with Fig. 9.2 are the following:



$$E_s(b) = \tau_{ss} E_s(t) , \quad (9.2.a)$$

$$E^-(b) = \tau_{sd} E_s(t) + \rho_{dd} E^+(b) , \quad (9.2.b)$$

$$E^+(b) = r_b [E_s(b) + E^-(b)] , \quad (9.2.c)$$

$$E_o(b) = r_t [E_s(b) + E^-(b)] , \quad (9.2.d)$$

$$E_o(t) = \rho_{so} E_s(t) + \tau_{do} E^+(b) + \tau_{oo} E_o(b) , \quad (9.2.e)$$

where (b) and (t) indicate the bottom and the top of the atmosphere. From equations (9.2.b) and (9.2.c) one finds the diffuse fluxes at the surface as

$$E^-(b) = E_s(t) (\tau_{sd} + \tau_{ss} r_b \rho_{dd}) / (1 - r_b \rho_{dd}) , \text{ and}$$

$$E^+(b) = E_s(t) (\tau_{ss} + \tau_{sd}) r_b / (1 - r_b \rho_{dd}) .$$

The total flux incident on the surface $E_{tot} = E_s(b) + E^-(b) = E_{sun} + E_{sky}$ is given by

$$E_{tot} = E_s(t) (\tau_{ss} + \tau_{sd}) / (1 - r_b \rho_{dd}) ,$$

and for $E_o(t)$, which is the radiance in the observer's direction multiplied by π , one finds

$$E_o(t) = \pi L_s = E_s(t) \left[\rho_{so} + \frac{\tau_{ss} + \tau_{sd}}{1 - r_b \rho_{dd}} (r_b \tau_{do} + r_t \tau_{oo}) \right] . \quad (9.3)$$

From comparison with Eq. (9.1) it follows that the path radiance can be found from

$$\pi L_p = E_s(t) \rho_{so} + E_{tot} r_b \tau_{do} , \text{ and that } T = \tau_{oo} .$$

$E_s(t)$ equals $E_s^o \cos \theta_s$, where E_s^o is the extraterrestrial solar (spectral) irradiance on a plane perpendicular to the sunrays, and θ_s is the solar zenith angle. Apart from the influence of the distance sun-earth during the year, E_s^o can be assumed constant. The ratio $\pi L_s / E_s(t)$ is the planetary reflectance r_p . The above description of the atmospheric effect is restricted to the case of a laterally homogeneous atmosphere over a surface for which both the target observed and the background act like Lambertian reflectors. As in general the background is not a homogeneous surface with a constant reflectance, r_b

should be considered an average reflectance over some neighbourhood around the target.

9.3 Extinction and scattering coefficients of the atmosphere

Because of the spherical shape of the particles, and if not, their random orientation, the atmosphere is an isotropic medium, which means that the interception coefficient β is independent of the direction of the incident radiation. In this case the extinction coefficients for the fluxes E_s , E_o and the couple (E^-, E^+) are given by

$$k = \beta / \mu_s ; \quad K = \beta / \mu_o ; \quad \kappa = 2 \beta \quad ,$$

$$\text{where } \mu_s = |\cos \theta_s| \quad \text{and } \mu_o = |\cos \theta_o| \quad .$$

The scattering of incident light in the atmosphere is primarily described by the scattering phase function $p(\delta)$ for the angular distribution, and by the single scattering albedo ω for the relative amount of scattering. In the atmosphere nearly all the intercepted light is scattered, so ω is usually close to one.

For the angular distribution one makes the distinction between Rayleigh-scattering by air molecules and Mie-scattering by aerosol particles with a size comparable to the wavelength of the radiation. For Rayleigh-scattering the phase function can be approximated as

$$p_R(\delta) = \frac{3}{4} (1 + \cos^2 \delta) \quad ,$$

where δ is the scattering angle, i.e. the angle between the incident and the exitent ray.

Mie-scattering depends on the particle size distribution, the wavelength and the complex index of refraction of the material (cf. Deirmendjian, 1969 and De Haan, 1987). Extensions of Mie-theory to non-spherical particles are discussed in De Haan (1987) and Stammes (1989). In general, the phase function of aerosols is highly peaked in the forward direction (the aureole region) and more or less oscillatory around the backward direction (the glory region), especially if the material is non-absorbing. Fig. 9.3 shows the Rayleigh phase function together with an example of Mie-scattering at a few

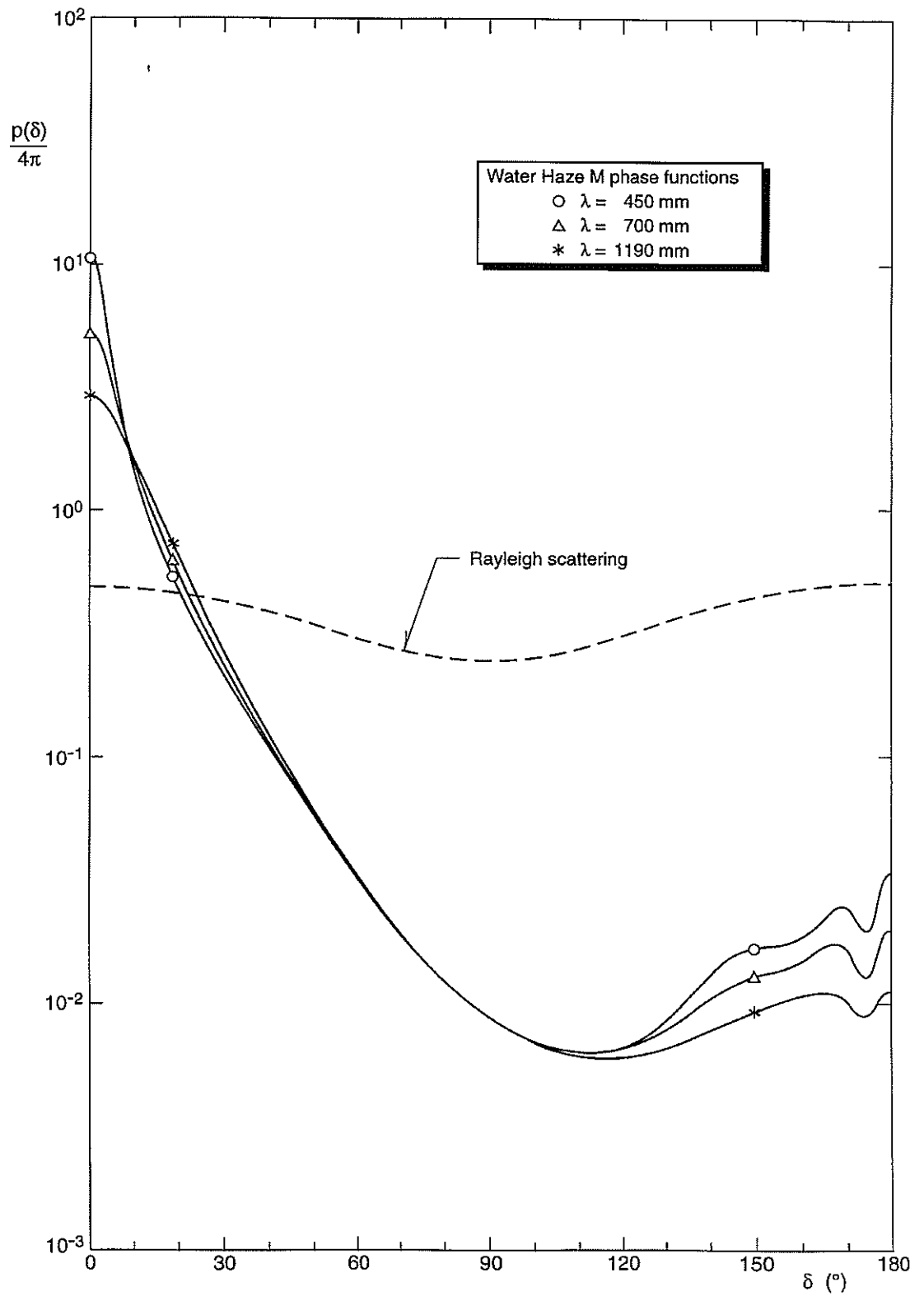


Fig. 9.3 Aerosol and Rayleigh scattering phase functions



wavelengths (water Haze M (maritime type) from Deirmendjian, 1969). From this it will be clear that the aerosol phase function $p_A(\delta)$ cannot be expressed by a simple function like the one for Rayleigh scattering.

The bidirectional scattering coefficient w is given by

$$w = \frac{\omega \beta p(\delta)}{4 \mu_s \mu_o} \quad , \text{ which holds for a single type of scattering.}$$

For a mixture of Rayleigh-scattering and aerosol-scattering it is formed by a linear combination as

$$w = \frac{\omega_R \beta_R p_R(\delta) + \omega_A \beta_A p_A(\delta)}{4 \mu_s \mu_o} \quad ,$$

where ω_R is the single scattering albedo for the Rayleigh case and can be taken equal to one. The subscripts R and A refer to Rayleigh and aerosol. Four-stream radiative transfer in the atmosphere (or any other scattering medium) is described by the matrix differential equation

$$\frac{d}{dz} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix} = \begin{bmatrix} -k & & & \\ s' & -(\kappa - \sigma') & \sigma & \\ -s & -\sigma & (\kappa - \sigma') & \\ -w & -v & -v' & K \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix} \quad , \quad (9.4)$$

in which z is the metrical depth, w is as given above and the extinction coefficients k , K , and κ for a mixture of Rayleigh and aerosol scattering are given by

$$k = (\beta_R + \beta_A) / \mu_s \quad , \quad K = (\beta_R + \beta_A) / \mu_o \quad , \quad \kappa = 2(\beta_R + \beta_A) \quad .$$

The remaining scattering coefficients are all found by integration of w over the upper and lower hemisphere, i.e.

$$\pi s' = \int_{-2\pi} w \mu_o d\Omega_o \quad , \quad \pi s = \int_{+2\pi} w \mu_o d\Omega_o \quad ,$$



$$\pi v' = \int_{-2\pi} w \mu_s d\Omega_s, \quad \pi v = \int_{+2\pi} w \mu_s d\Omega_s,$$

$$\pi \sigma' = \int_{+2\pi} s' \mu_s d\Omega_s = \int_{+2\pi} v' \mu_o d\Omega_o, \quad \pi \sigma = \int_{+2\pi} s \mu_s d\Omega_s = \int_{+2\pi} v \mu_o d\Omega_o,$$

where $(-2\pi, +2\pi)$ indicates integration over the (lower, upper) hemisphere.

For Rayleigh scattering the integration of the phase function over one hemisphere gives always as a result 2π , half of the spherical integral. For the aerosol phase function this is not the case. Integration over the backward hemisphere (i.e. the upper hemisphere if $\mu_s = 1$ or $\theta_s = 0$) gives the so-called backscattering efficiency η_A as

$$\eta_A = \frac{1}{4\pi} \int_0^{2\pi} \int_0^1 P(\mu_s = 1, \mu_o, \varphi_o) d\mu_o d\varphi_o.$$

For most aerosol types η_A is of the order of 0.05, so 95 percent is scattered into the forward hemisphere.

Approximated values of the above scattering coefficients can now be found by assuming that always a fraction η_A of the incident flux is scattered into the hemisphere of incidence and a fraction $1 - \eta_A$ into the opposite hemisphere. This gives

$$\begin{aligned} s' &= \left[\frac{1}{2} \beta_R + \omega_A (1 - \eta_A) \beta_A \right] / \mu_s, & s &= \left[\frac{1}{2} \beta_R + \omega_A \eta_A \beta_A \right] / \mu_s, \\ v' &= \left[\frac{1}{2} \beta_R + \omega_A (1 - \eta_A) \beta_A \right] / \mu_o, & v &= \left[\frac{1}{2} \beta_R + \omega_A \eta_A \beta_A \right] / \mu_o, \\ \sigma' &= \beta_R + 2 \omega_A (1 - \eta_A) \beta_A, & \sigma &= \beta_R + 2 \omega_A \eta_A \beta_A, \end{aligned}$$

in which ω_R was assumed to be equal to one.

The combination $(\kappa - \sigma')$ in Eq. (9.4) is called the attenuation coefficient a , and is given by



$$a = \beta_R + 2 [1 - \omega_A (1 - \eta_A)] \beta_A \quad .$$

For the solution of Eq. (9.4) it is first assumed that the result will not depend much on the profiles of β_R and β_A as a function of the height h in the atmosphere. This is equivalent to assuming that the atmosphere forms a homogeneous mixture of Rayleigh and aerosol scattering, with associate optical thicknesses b_R and b_A , respectively. These are defined by

$$b_R = \int_0^{\infty} \beta_R(h) dh \quad \text{and} \quad b_A = \int_0^{\infty} \beta_A(h) dh \quad .$$

Assigning an arbitrary height H to the homogeneous atmospheric layer now gives the equivalent average interception coefficients β_R' and β_A' as

$$\beta_R' = b_R/H \quad \text{and} \quad \beta_A' = b_A/H \quad .$$

These can be used instead of β_R and β_A for the definition of the extinction and scattering coefficients. However, as the actual height is irrelevant for the solution of Eq. (9.4), one can just as well take $H = 1$, so that $\beta_R' = b_R$ and $\beta_A' = b_A$. When Eq. (9.4) is written as

$$\frac{d}{dz} E = M' E \quad ,$$

where M' is the matrix of coefficients with use of β_R' and β_A' ,

the concept of relative optical height is introduced as follows:

$$\text{Writing } dE = M' E dz = (1/H) M E dz = M E \frac{dz}{H} = M E dx \quad ,$$

where M is the matrix of coefficients for $H = 1$, leads to the replacement of the z co-ordinate by the relative optical height $x = z/H$. The range of x is also arbitrary, but it appears convenient if one takes $x = 0$ for the top and $x = -1$ for the bottom of the layer. This means that $x = (z-H)/H$. One can now write

$$\frac{d}{dx} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix} = \begin{bmatrix} k & & & \\ -s' & a & -\sigma & \\ s & \sigma & -a & \\ w & v & v' & -K \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o \end{bmatrix}, \quad (9.5)$$

where x runs from -1 to 0 , and

$$\begin{aligned} k &= (b_R + b_A) / \mu_s ; \quad K = (b_R + b_A) / \mu_o ; \\ w &= [b_R p_R(\delta) + \omega_A b_A p_A(\delta)] / (4 \mu_s \mu_o) ; \\ s' &= [1/2 b_R + \omega_A (1 - \eta_A) b_A] / \mu_s ; \quad s = [1/2 b_R + \omega_A \eta_A b_R] / \mu_s ; \\ v' &= [1/2 b_R + \omega_A (1 - \eta_A) b_A] / \mu_o ; \quad v = [1/2 b_R + \omega_A \eta_A b_R] / \mu_o ; \\ a &= b_R + 2 [1 - \omega_A (1 - \eta_A)] b_A ; \quad \sigma = b_R + 2 \omega_A \eta_A b_A . \end{aligned}$$

To the coefficients k , K and a can be added contributions due to gaseous absorption (for instance by water vapour) in the atmosphere. If the optical thickness associated with this absorption is called b_G , then these extra contributions are equal to b_G / μ_s , b_G / μ_o and $2 b_G$, respectively.

The solution of Eq. (9.5) can be expressed in matrix-vector form as

$$\begin{bmatrix} E_s(-1) \\ E^-(-1) \\ E^+(0) \\ E_o(0) \end{bmatrix} = \begin{bmatrix} \tau_{ss} & & & \\ \tau_{sd} & \tau_{dd} & \rho_{dd} & \\ \rho_{sd} & \rho_{dd} & \tau_{dd} & \\ \rho_{so} & \rho_{do} & \tau_{do} & \tau_{oo} \end{bmatrix} \begin{bmatrix} E_s(0) \\ E^-(0) \\ E^+(-1) \\ E_o(-1) \end{bmatrix}, \quad (9.6)$$

in which the nine reflectance and transmittance factors are functions of the extinction and scattering coefficients as defined above. In Verhoef (1985a) it was shown that these functions are rather simple, for instance



$$\tau_{ss} = e^{-k} \quad , \quad \tau_{oo} = e^{-K} \quad ,$$

$$\rho_{dd} = \frac{e^m - e^{-m}}{r_\infty^{-1} e^m - r_\infty e^{-m}} \quad , \quad \tau_{dd} = \frac{r_\infty^{-1} - r_\infty}{r_\infty^{-1} e^m - r_\infty e^{-m}} \quad ,$$

$$\text{where } m = \sqrt{a^2 - \sigma^2} \quad \text{and} \quad r_\infty = (a - m)/\sigma \quad .$$

Here, ρ_{dd} and τ_{dd} are equal to corresponding expressions of the Kubelka-Munk two-stream theory.

A special case, but not uncommon for atmospheric scattering, is the one encountered if $\omega_A = 1$ and $b_G = 0$ (no absorption at all). In this case $a = \sigma$, so $m = 0$, and ρ_{dd} and τ_{dd} become indeterminate if calculated according to the given expressions. This singularity can be removed if the limits for $m \rightarrow 0$ are taken, which gives

$$\rho_{dd} \underset{m \rightarrow 0}{=} \frac{\sigma}{\sigma + 1} \quad \text{and} \quad \tau_{dd} \underset{m \rightarrow 0}{=} \frac{1}{\sigma + 1} \quad .$$

The sum of both equals one, which is consistent with the absence of any absorption in the atmosphere.

9.4 Model implementation aspects

Of the nine reflectance and transmittance factors of Eq. (9.6), only six are needed for the determination of the atmospheric effect, namely the reflectance factors ρ_{so} and ρ_{dd} , and the transmittance factors τ_{ss} , τ_{sd} , τ_{do} and τ_{oo} , as can be seen in Fig. 9.2. These parameters describe the effects of Rayleigh scattering, scattering by aerosols and, possibly, absorption by gases like water vapour. However, absorption by ozone gas takes place mainly at altitudes of 20 to 25 km, well above the layer where the above mentioned processes are concentrated. Therefore, it is better to incorporate ozone absorption into the model by adding a separate layer at the top of the tropospheric layer. In this "ozone layer" only absorption is supposed to take place, no scattering. The transmittance factors associated with absorption by



ozone can be symbolized as T_{sO_3} and T_{oO_3} and are equal to

$$T_{sO_3} = e^{-b_{O_3}/\mu_s} \quad \text{and} \quad T_{oO_3} = e^{-b_{O_3}/\mu_o} ,$$

where b_{O_3} is the optical thickness due to ozone absorption. The reflectance and transmittance factors of the complete atmosphere should now be modified as

$$\begin{aligned} \rho_{so}^* &= T_{sO_3} \rho_{so} T_{oO_3} ; & \tau_{ss}^* &= T_{sO_3} \tau_{ss} ; & \tau_{sd}^* &= T_{sO_3} \tau_{sd} \\ \tau_{do}^* &= \tau_{do} T_{oO_3} ; & \tau_{oo}^* &= \tau_{oo} T_{oO_3} \end{aligned}$$

As ρ_{dd} in the description of the atmospheric effect only plays the role of a spherical albedo at the bottom of the atmosphere, it does not need to be modified.

Values of b_{O_3} at several wavelengths can be found in the literature. Apart from the well-known strong absorption of ultraviolet light by ozone, in the visible some additional absorption takes place, with a maximum at about 600 nm, where b_{O_3} is of the order of 0.04 (cf. Elterman, 1970).

The optical thickness associated with Rayleigh scattering, b_R , depends strongly on the wavelength. According to Elterman (1970) it can be approximated by

$$b_R = 0.0987 \left[\frac{\lambda}{550} \right]^{-4.06}, \quad \text{where } \lambda \text{ is the wavelength in nm.}$$

This means that halving the wavelength gives a more than 16-fold increase of the Rayleigh scattering optical thickness. The relationship given above holds for standard air pressure and temperature at sea level, but corrections for actual conditions can be carried out easily.

For aerosol scattering the dependence of b_A on wavelength is much weaker. If expressed as $b_A = \beta \lambda^\alpha$, where β and α are constants, the value of α , called the Ångström coefficient, is usually between -0.6 and -1.3 . However, this only applies to particle size distributions of a special type, namely the Junge-distribution, which is of the power law type. For the modified gamma distributions introduced by Deirmendjian (1969), the dependence on wavelength is different, although still smooth. Plotted on a



log-log scale, Deirmendjian's curves are convex, with a negative slope which becomes more negative with increasing wavelength. This means that in this case α is not a constant, but itself a function of the wavelength. For the atmospheric correction model developed at NLR by the author, one of Deirmendjian's tabulated aerosol phase functions, namely the one for water Haze M (maritime type), was selected as a prototype for the representation of aerosol scattering behaviour in general. This function is tabulated at 34 values of the scattering angle δ and at a number of wavelengths in the range from 450 to 2250 nm. For the computation of $p_A(\delta)$ at arbitrary δ and λ , a cubic spline interpolation is carried out with respect to δ , and linear interpolation with respect to λ . The single scattering albedo ω_A for this type of aerosol is practically equal to one, but in the computer program smaller values of ω_A are allowed if so desired.

Water vapour in the atmosphere has absorption bands mainly in the infrared part of the spectrum, for instance at 930, 1150, 1400 and 1900 nm. Most optical remote sensing instruments avoid these bands, but the Landsat MSS scanner and the NOAA-AVHRR (advanced very high resolution radiometer) both have a spectral band that includes the absorption peak at 930 nm. The optical thickness associated with this is of the order of 0.1 (Saunders, 1988), but, like temperature and humidity, its variability is high, both spatially and temporarily. For instruments like the Landsat Thematic Mapper and the SPOT HRV the influence of water vapour absorption can probably be ignored in most cases.

Summarized, the greatest uncertainties in the atmospheric model are associated with the aerosol properties b_A , ω_A and $p_A(\delta)$, as these are highly variable or difficult to measure. Of these, b_A is the most important quantity because of its large influence on the extinction and scattering coefficients. The other two quantities are less variable than b_A , since ω_A will mostly be rather close to one and $p_A(\delta)$ is roughly similar for different types of aerosol, especially in the range of δ involved in remote sensing missions ($\delta > \pi/2$). Therefore it appears not unreasonable to adopt representative examples of ω_A and $p_A(\delta)$, so that the only unknown left would be b_A .



9.5 Estimation of aerosol optical thickness

Several techniques can be applied to estimate the aerosol optical thickness b_A . The first, and the least reliable one, makes use of the parameter meteorological visual range, or visibility, which is defined as the horizontal distance at sea level over which the contrast at a wavelength of 550 nm is reduced to 2 percent of the one at zero distance. As this contrast reduction equals the direct horizontal transmittance T_{hor} , given by

$$T_{hor} = e^{-\beta(0)d}, \text{ where } \beta(0) \text{ is the interception coefficient at 550 nm in } \text{km}^{-1} \text{ at sea level and } d \text{ is the distance in km, one may put for } T_{hor} = 0.02:$$

$$0.02 = e^{-\beta(0)V}, \text{ where } V \text{ is the visibility in km, or}$$

$$\beta(0) = (\ln 50)/V.$$

As $\beta(0)$ is the sum of the contributions due to Rayleigh scattering and aerosol scattering, one may write

$$\beta_R(0) + \beta_A(0) = (\ln 50)/V, \text{ or } \beta_A(0) = (\ln 50)/V - \beta_R(0).$$

From the literature, $\beta_R(0)$ at 550 nm is known to be equal to 0.0116, leading to $V = 337$ km if $\beta_A(0) = 0$. Such a high visibility is never found in reality at sea level. For cloudless atmospheres, a more realistic range for V is from 5 to 40 km. In that case $\beta_A(0)$ ranges from 0.086 at $V = 40$ km to 0.771 at $V = 5$ km.

A different measure of the state of the atmosphere is the turbidity factor T , which is defined as $T = (b_R + b_A)/b_R$, so it refers to the entire atmosphere, not to the situation at sea level, where V is based on.

Use of the visibility for estimation of b_A can only be carried out if it is known how β_A depends on the height in the atmosphere if $\beta_A(0)$ is given. A simple model of this profile was discussed in Sturm (1981) and consists of the following equations:

$$\beta_A(h) = \begin{cases} \beta_A(0) e^{-h/H_1} & \text{if } h \leq 5.5 \text{ km} \\ \beta_{A5.5} & \text{if } 5.5 \text{ km} < h < 18 \text{ km} \\ \beta_{A5.5} e^{(18-h)/H_2} & \text{if } h \geq 18 \text{ km} \end{cases}$$

Together with that for Rayleigh scattering, this profile is illustrated in Fig. 9.4. For Rayleigh scattering $\beta_R(h) = \beta_R(0) e^{-h/H_0}$, where $H_0 = 8.5155$ km.

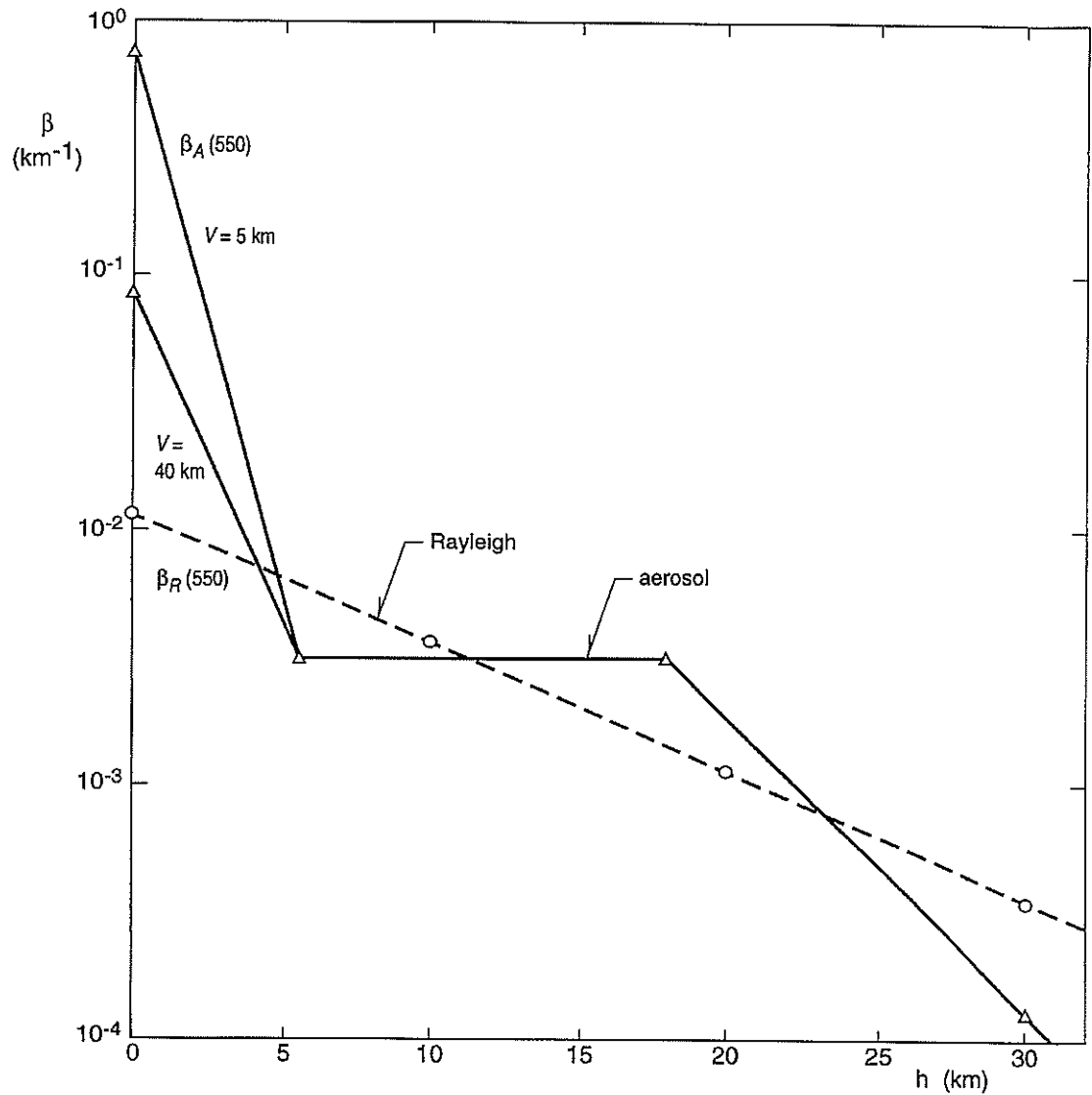


Fig. 9.4 Height profiles of the interception coefficients for aerosol and Rayleigh scattering, according to Sturm (1981)

Here H_0 , H_1 and H_2 are so-called scale heights. Of these, H_0 and H_2 are considered constant, but H_1 is related to $\beta_A(0)$ by the requirement that $\beta_A(h)$ is continuous at $h = 5.5$ km, giving $\beta_A(0) e^{-5.5/H_1} = \beta_{A5.5}$, or

$$H_1 = 5.5 / \ln [\beta_A(0) / \beta_{A5.5}] \quad .$$

In Sturm's profile $H_2 = 3.748$ km and $\beta_{A5.5}$ at $\lambda = 550$ nm equals $0.0030765 \text{ km}^{-1}$. In that case the aerosol optical thickness of the layer above $h = 5.5$ km equals $\beta_{A5.5} (18 - 5.5 + H_2) = 0.05$. For the layer below $h = 5.5$ km one obtains an optical thickness of $[\beta_A(0) - \beta_{A5.5}] H_1$, so that the total optical thickness at $\lambda = 550$ nm is given by

$$b_A = [\beta_A(0) - \beta_{A5.5}] H_1 + 0.05 \quad .$$

For visibilities V of 5 km and 40 km this gives $b_A = 0.815$ and $b_A = 0.187$, respectively. Corresponding values of the turbidity factor T are 9.26 and 2.89.

From the example above it appears that b_A at $\lambda = 550$ nm can be estimated if the visibility V is given and if one can be confident that the actual profile of β_A more or less matches the modelled profile as a function of the height. As shown by the example, the influence of the layer above $h = 5.5$ km is only small, so the greatest errors are expected to be associated with differences between actual and modelled profile in the lowest 5.5 km of the atmosphere. There is one particular situation in which the actual profile can be very different from the modelled one, namely in the case of an inversion layer in the atmosphere. In that case the normal decrease of the air temperature with height is interrupted by a layer in which it is constant or increases. In such a situation a haze layer can develop at some height, while the atmosphere at the surface may still be relatively clear. It is obvious that in this case the visibility at ground level can only be a very poor indicator of the total aerosol optical thickness.

Other drawbacks of this technique are that visibility often is not measured but only visually estimated by a human observer, and that extrapolation of b_A to wavelengths other than $\lambda = 550$ nm is questionable.

The second technique uses model inversion, i.e. the estimation of model parameters from measurements, in order to estimate b_A . Some possibilities for this can be illustrated by means of Fig. 9.5, which shows the behaviour of some model output quantities as a function of b_A in the range from 0.0 to 1.0. All quantities shown are relative to the extraterrestrial solar irradiance, E_s^o , and the influence of the aerosol single scattering albedo ω_A on the results



is demonstrated by showing results for $\omega_A = 1.0$ and $\omega_A = 0.9$. As appears from Fig. 9.5, measurement of E_{sun} gives the best estimate of b_A , since it is independent of ω_A . However, the value of E_s^o in the spectral band over which E_{sun} is measured must be known and E_{sun} must be measured with a well-calibrated instrument, which sometimes can be problematic, as the usual absolute calibration accuracy is only of the order of ten percent. Measurement of both E_{sun} and E_{sky} , for instance by use of a Guzzi spectroradiometer (Veugen & Van Stokkom, 1985), can give good estimates of both b_A and ω_A , provided the calibration is accurate. This instrument has a rotating band

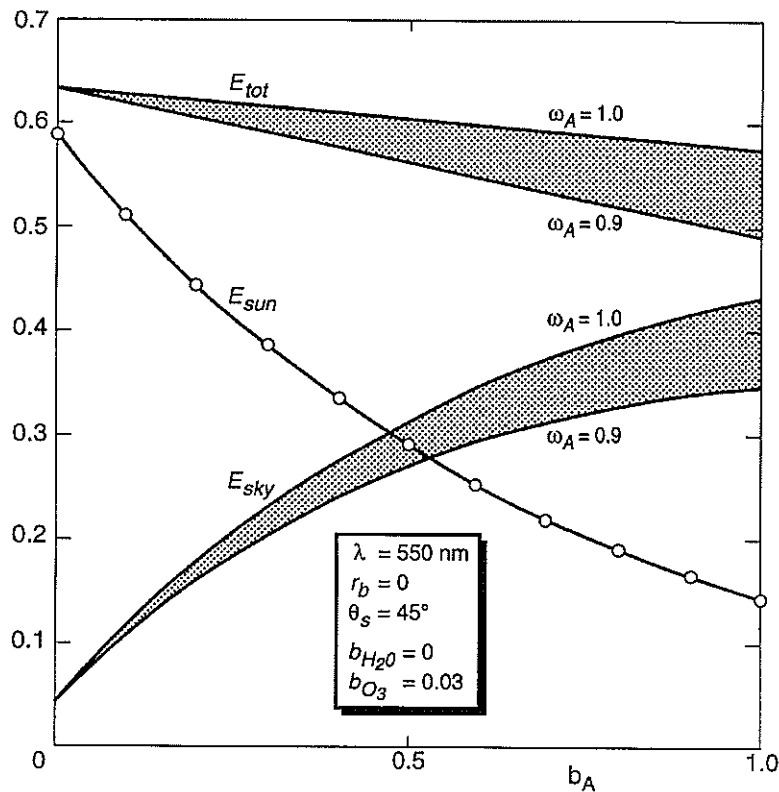


Fig. 9.5 Solar, sky and total irradiance at ground level (relative to E_s^o) as a function of b_A and ω_A at $\lambda = 550 \text{ nm}$

which periodically blocks the sunlight from entering the detector, so a continuous measurement of E_{tot} and E_{sky} during the day is possible. If the calibration is questionable, then one can still get fairly accurate results for b_A by taking the ratio E_{sky}/E_{tot} , since in that case the calibration error is cancelled, and the influence of ω_A on this ratio is much smaller than on E_{sky} and E_{tot} separately.

Under favourable circumstances the ratio E_{sky}/E_{tot} can be measured from an



image, for instance when a small cumulus cloud throws its shadow on a large homogeneous piece of land. The digital numbers of this object in the shadow and in the sunlit part can then be used for an estimate of the ratio E_{sky}/E_{tot} .

Another quantity that can be measured from an image is the atmospheric planetary reflectance $r_p(0)$, which is the planetary reflectance for zero ground reflectance. Under the condition that an object and its surroundings have a reflectance close to zero (for instance coniferous forest in the visible blue and red, or clear water in the near infrared), it may be assumed that the digital number DN for such a dark object can solely be attributed to atmospheric reflectance, so that after calibration of DN in units of reflectance an estimate of $r_p(0)$ is found. Provided the atmosphere over the scene is homogeneous, the lowest DN values of the scene can be associated with these dark objects, and the method based on this idea is therefore called the "darkest pixel" method.

The potentials of both quantities E_{sky}/E_{tot} and $r_p(0)$ for estimation of b_A are illustrated in Figures 9.6 and 9.7, respectively. Here, these quantities are shown as a function of b_A at 550 nm, under the assumption that $\alpha = -1.0$. Results are plotted for two wavelengths, namely $\lambda = 450$ nm (blue) and $\lambda = 700$ nm (red), and two values of ω_A , namely 0.9 and 1.0.

From Fig. 9.6 one may conclude that measurement of the ratio E_{sky}/E_{tot} can give a good estimate of the aerosol optical thickness b_A and that the influence of ω_A on this estimate is only moderate. Measurement of this quantity at two or more wavelengths can be used to estimate the Ångström-coefficient α as well.

As appears from Fig. 9.7, the relation between the atmospheric planetary reflectance $r_p(0)$ and b_A can also be used to estimate b_A , but the influence of ω_A on this relationship is considerable, especially at $\lambda = 450$ nm. Therefore it can be concluded that the "darkest pixel method" cannot give good estimates of b_A , unless one has reason to believe that ω_A is very close to one, for instance when the aerosol is known to be of oceanic origin. Continental and urban aerosols usually contain more absorbing materials like soot and dust, and in that case ω_A can be much smaller. The strong wavelength-dependence of $r_p(0)$ suggests that the Ångström-coefficient can still be estimated fairly well from measurements of $r_p(0)$ at two or more wavelengths.

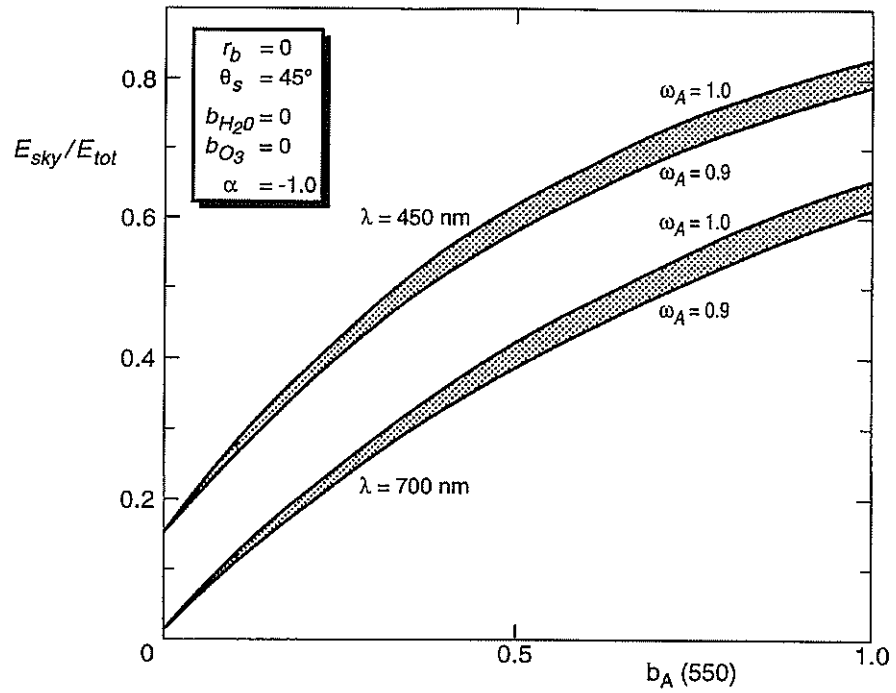


Fig. 9.6 The ratio E_{sky} / E_{tot} at $\lambda = 450$ nm and 700 nm as a function of b_A at $\lambda = 550$ nm for $\alpha = -1$ and $\omega_A = 0.9$ to 1.0

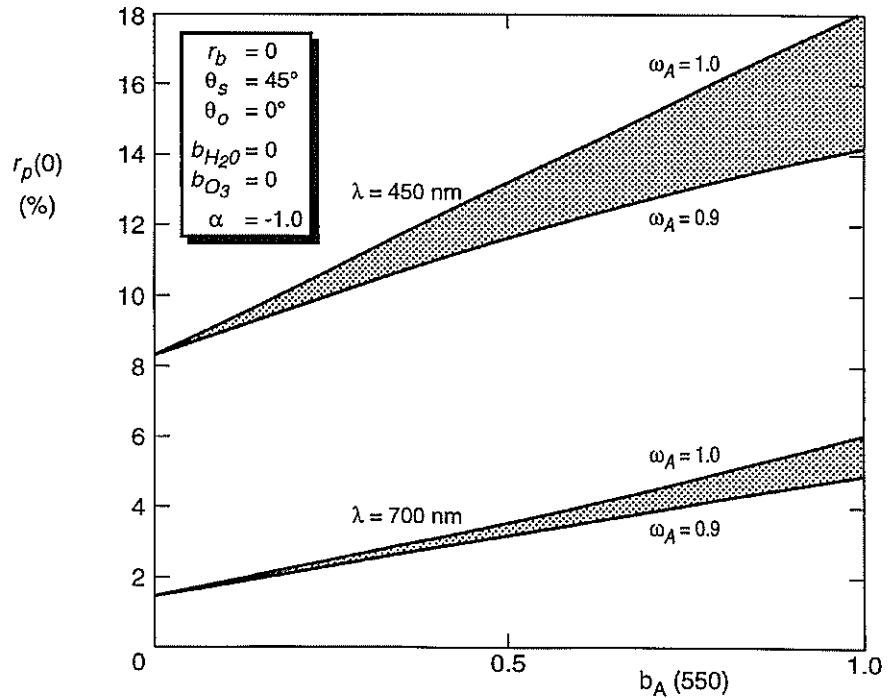


Fig. 9.7 The atmospheric planetary reflectance $r_p(0)$ under the same conditions as in Fig. 9.6



9.6 Atmospheric correction of Landsat Thematic Mapper images

On the basis of the material discussed in the previous sections a procedure for the correction of atmospheric effects and calibration in units of ground reflectance has been developed for Landsat Thematic Mapper images. The Thematic Mapper instrument is an opto-mechanical multispectral scanner with 30 m ground resolution and has six spectral bands in the optical region and one in the thermal infrared (with 120 m ground resolution). The six optical spectral bands are called TM1 to TM5 plus TM7 (TM6 is the thermal infrared band) and are centred at wavelengths of 485, 560, 660, 830, 1650 and 2215 nm (visible, near and middle infrared), respectively.

Digital images acquired by the Landsat Thematic Mapper can be ordered from receiving stations in the USA, Europe and other locations.

Apart from the digital images, each data set contains also an extensive amount of auxiliary information, such as the scene location, time of acquisition, solar elevation angle and calibration constants. By means of the calibration data for each spectral band it is possible to calculate the detected radiances L_s from the digital number DN of an image pixel as

$$L_s = A_0 + A_1 * DN \quad , \quad (9.7)$$

where A_0 = offset ,

A_1 = gain factor .

The radiance detected at the satellite L_s can be related to the planetary reflectance r_p by

$$\pi L_s = r_p E_s^o \cos \theta_s / d^2 \quad , \quad (9.8)$$

where E_s^o is the extraterrestrial solar irradiance in the associate spectral band at a sun-earth distance of 1 Astronomical Unit (AU), and d is the actual distance in AU. The distance d is season-dependent, with a minimum on 3 January of 0.983 and a maximum on 2 July of 1.017, so that d^2 varies by about 7 percent at maximum. The average of d over a year is, by definition, equal to one.

The planetary reflectance of an image pixel with digital number DN in a spectral band can be calculated by combining both equations, which gives



$$r_p = \frac{\pi (A_0 + A_1 * DN) d^2}{E_s^o \cos \theta_s} \quad (9.9)$$

It is important to note that E_s^o should be given in units which are compatible with those of A_0 and A_1 . In this respect, the distributor of Landsat data in the USA, Eosat, specifies A_0 and A_1 in $\text{mW}/(\text{cm}^2 \text{ sr } \mu\text{m})$, which means that the calibration constants are spectral radiances. In this case the compatible unit for E_s^o is $\text{mW}/(\text{cm}^2 \mu\text{m})$. On the other hand, the European distributor Eurimage applies units of $\text{W}/(\text{m}^2 \text{ sr})$, which refers to so-called in-band radiance, the integral of the spectral radiance over the width of the spectral band. In that case the compatible unit for E_s^o is W/m^2 .

Besides the above complications, the user of Landsat TM data is also confronted with the fact that values of E_s^o for the TM bands are not given by the distributors, but have to be found in the specialist literature (cf. Markham & Barker, 1987). A good review of the difficulties associated with the calibration of Landsat TM data is given in Epema (1990). As a final remark it can be stated that much of the confusion could be avoided if the distributors would supply the users with alternative calibration constants B_0 and B_1 defined as $B_0 = \pi A_0 / E_s^o$ and $B_1 = \pi A_1 / E_s^o$. In that case,

$$r_p = (B_0 + B_1 * DN) d^2 / \cos \theta_s \quad (9.10)$$

and there can be no misunderstanding about the units of B_0 and B_1 , as they are dimensionless. In addition, the user would not have to consult other sources of information, as was the case with E_s^o .

According to Eq. (9.3) of section 9.2, the planetary reflectance r_p is given by

$$r_p = \rho_{so} + \frac{\tau_{ss} + \tau_{sd}}{1 - r_b \rho_{dd}} (r_b \tau_{do} + r_t \tau_{oo}) \quad (9.11)$$

where the atmospheric parameters ρ_{so} , ρ_{dd} , τ_{ss} , τ_{sd} , τ_{do} and τ_{oo} can be obtained from model calculations, r_b is the background reflectance and r_t is the target reflectance. Since the background reflectance r_b is not known *a priori*, in a first approximation it may be assumed that r_b equals r_t . Of course this can only be correct if the target is large enough. In that case,



$$r_p = \rho_{so} + \frac{T_1 T_2 r_t}{1 - r_t \rho_{dd}} \quad , \text{ where } T_1 = \tau_{ss} + \tau_{sd} \text{ and } T_2 = \tau_{oo} + \tau_{do} \quad .$$

Solution of r_t from this equation gives

$$r_t = \frac{r_p - \rho_{so}}{T_1 T_2 + (r_p - \rho_{so}) \rho_{dd}} \quad . \quad (9.12)$$

So, for this simple atmospheric correction method it is only necessary to calculate r_p from DN using the calibration data, and to know the atmospheric parameters ρ_{so} , ρ_{dd} and the product $T_1 T_2$ for each spectral band. As explained in the previous chapter, the most unknown quantity for the determination of the atmospheric parameters is the aerosol optical thickness b_A . However, it was shown that b_A can be estimated from measured data of the ratio E_{sky}/E_{tot} or from values of the darkest pixel planetary reflectance $r_p(0)$ extracted from the image. In both cases model inversion is applied to estimate b_A . This is done by means of iteration, i.e. the input value b_A is varied until the calculated output quantity matches the measured value. If this technique is applied for two or more wavelengths, then the wavelength-dependence of b_A can be determined by means of the assumed relationship $b_A(\lambda) = \beta \lambda^\alpha$, or $\log b_A(\lambda) = \log \beta + \alpha \log \lambda$. Linear regression of $\log b_A(\lambda)$ against $\log \lambda$ thus provides least squares estimates of the parameters β and α . Subsequently, β and α can be applied in order to estimate $b_A(\lambda)$ in all the Thematic Mapper bands in the optical region, after which also ρ_{so} , ρ_{dd} and $T_1 T_2$ can be calculated by means of the atmospheric model in all spectral bands.

With respect to the darkest pixel method it should be mentioned that the linear regression $\log b_A - \log \lambda$ leads to a straight line (best fit) around which the model-inverted values b_A^* are scattered if more than two wavelengths were used, so that some points will lie above the regression line and some will lie under it. For the wavelengths at which b_A^* is under the regression line, application of the best fit value b_A in the calculation of the atmospheric parameters ρ_{so} , ρ_{dd} and $T_1 T_2$ will result in an over-estimation of the atmospheric effect. Especially, ρ_{so} will be greater than the darkest pixel planetary reflectance $r_p(0)$ in that case, so that for the darkest pixels negative values of the target reflectance r_t would be computed. In order to avoid this,



the regression line is lowered parallel to itself until it goes through the point having the largest difference between $\log b_A$ and $\log b_A^*$. In this way the Ångström coefficient α remains the same, but β is lowered to a new value β' . Next, the relationship $b_A(\lambda) = \beta' \lambda^\alpha$ is used for the estimation of ρ_{so} , ρ_{dd} and $T_1 T_2$ in all reflective TM-bands by means of the model.

The actual correction of an image by means of equations (9.9) and (9.12) is carried out by means of a look-up table (LUT) for each spectral band. As Landsat TM images are 8 bit per band, each input pixel DN is in the range 0-255, so each LUT has 256 entries. If the corrected image is also encoded in 8 bits per band, and the corrected digital number is called DN' , then the correction is carried out by applying the operation $DN' = \text{LUT}(DN)$ to each pixel in the image for each spectral band. As DN' represents a reflectance value in the range 0.0 - 1.0, it is necessary to specify a scale factor s which relates reflectance to digital number. This scale factor should be smaller than 256 in order to prevent overflow.

The entire procedure for atmospheric correction of a Landsat TM image is summarized below for the darkest pixel method:

1. Enter date $\rightarrow d^2$
 Enter θ_s
 Enter TM-calibration data file name
2. For two or more TM bands:
 Enter darkest pixel DN
 $DN \rightarrow r_p$ (calibration)
 Iterate b_A^* until $\rho_{so}(\text{model}) = r_p$
3. Linear regression $\log b_A^*$ vs. $\log \lambda \rightarrow \log b_A = \log \beta + \alpha \log \lambda$
 Lower regression line $\rightarrow \log b_A = \log \beta' + \alpha \log \lambda$
4. For all TM bands:
 $b_A = \beta' \lambda^\alpha$
 model $\rightarrow \rho_{so}, \rho_{dd}, T_1 T_2$
5. Enter scale factor s
6. For all TM bands:
 For $DN = 0, 1, \dots, 255$:
 $DN \rightarrow r_p$ (calibration)
 $r_p \rightarrow r_t$ (atmospheric correction)
 $\text{LUT}(DN) = s * r_t$

For all pixels:

$$DN' = \text{LUT}(DN)$$

7. End

For the method based on measurements of E_{sky}/E_{tot} the procedure is similar, except that step 2 consists of:

2. For two or more wavelengths:

Enter measured E_{sky}/E_{tot}

Iterate b_A^* until $E_{sky}/E_{tot}(\text{model}) = E_{sky}/E_{tot}(\text{measured})$

Also, the lowering of the regression line in step 3 is omitted in this case.

For both methods of correction the computation time is negligible, mainly because the model is simple, the iteration converges rapidly and the look-up table operation used in the correction can be carried out very efficiently.

9.7 Validation results

Although in practice a true validation of an atmospheric model is a very difficult task because of the great number of parameters that would have to be measured, and no specific attempts have been initiated in this direction, the model and the correction methods based on it have been used in several projects, in all cases with the aim to improve the spectral characterization of objects on the ground and to facilitate multitemporal comparison.

In one project, described by Epema (1992), also *in situ* ground measurements of the reflectance were available, so that the performance of the correction method could be tested. In this case the method based on the ratio E_{sky}/E_{tot} was used in Tunisia, where this ratio was measured by shadowing a reference panel. The main conclusions were that for a Landsat TM image of April 1988 the correspondence between ground reflectances derived from the image and measured values was good (maximum relative error 10 percent), but for an image of December 1987 larger errors were observed. However, these could be attributed to water vapour absorption in TM bands 4, 5 and 7. After incorporating this water vapour absorption into the model the results improved significantly.

In another project (Verhoef, 1990) the darkest pixel method has been tested for images of the Flevoland area in The Netherlands, acquired in the summer



of 1986. Some results of this exercise are discussed below.

Tables 9.1 and 9.2 summarize the numerical results of the model inversion and the computed correction constants for Landsat TM images of 16 June and 3 August 1986, respectively. In both cases, four TM bands were used for the estimation of b_A from darkest pixel digital numbers. Of these, TM1 (blue) and TM4 (near infrared) are the most reliable ones, since the darkest pixels in these bands most likely refer to objects having a reflectance very close to zero, such as coniferous forest in TM1 and clear water in TM4, which both are present in the images. In TM2 (green) and TM3 (red) the darkest objects probably still have a reflectance of the order of one percent, and this might lead to overestimation of b_A in these bands, as this small reflectance would wrongly be attributed to the atmospheric effect.

Therefore, the procedure was modified by allowing the specification of a small reflectance for the darkest object. In the tables these are indicated by r_i for the different spectral bands. The results of the linear regression $\log b_A$ vs. $\log \lambda$ are given by the parameters α and β , the squared correlation coefficient R^2 and the root mean square error (RMSE) of b_A . For both dates a very good fit was found, as expressed by the high values of R^2 and the small values of the RMSE. Also, the Ångström coefficient α is in the expected range of -0.6 to -1.3 .

Comparing the results for both dates, there appears to be a large difference in the atmospheric turbidity, as evidenced by the difference in the aerosol optical thickness b_A , but this does not lead to very great differences in the correction constants ρ_{so} , $T_1 T_2$ and ρ_{dd} . The reason for this is probably the fact that at shorter wavelengths Rayleigh scattering, which is virtually constant, tends to dominate the atmospheric effect, and that at the longer wavelengths the atmospheric effect is small anyway, except for possible effects due to water vapour absorption, as found by Epema.

From the two atmospherically corrected and calibrated images, TM-derived spectral reflectance 'signatures' have been extracted for a number of different objects in order to evaluate the performance of the correction. The results are presented in Figures 9.8 and 9.9. In both Figures the solid lines refer to the 16 June image and the dashed lines to the 3 August image. In Fig. 9.8 also results of field spectrometer measurements for grass and sugar beet, as obtained during the NIWARS programme in 1973 (Bunnik, 1978), have been included for comparison. The TM-derived grass spectrum of 16 June appears to be very similar to the measured NIWARS spectrum of 1 August 1973. This difference in date is of no significance, however, as grass

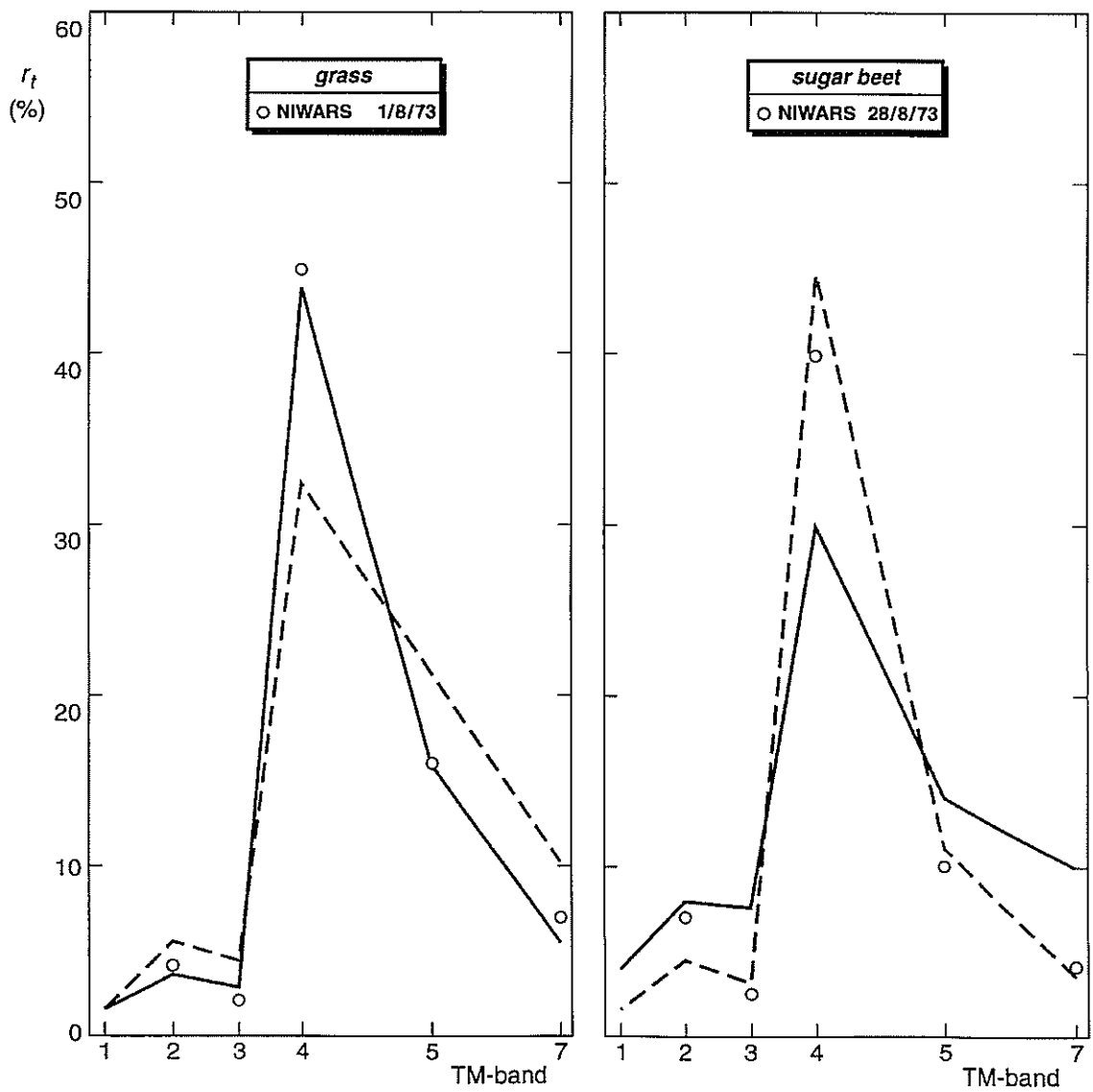


Fig. 9.8 *TM-derived spectral reflectance for grass and sugar beet compared with NIWARS field spectrometer measurements in 1973 (Bunnik, 1978)*

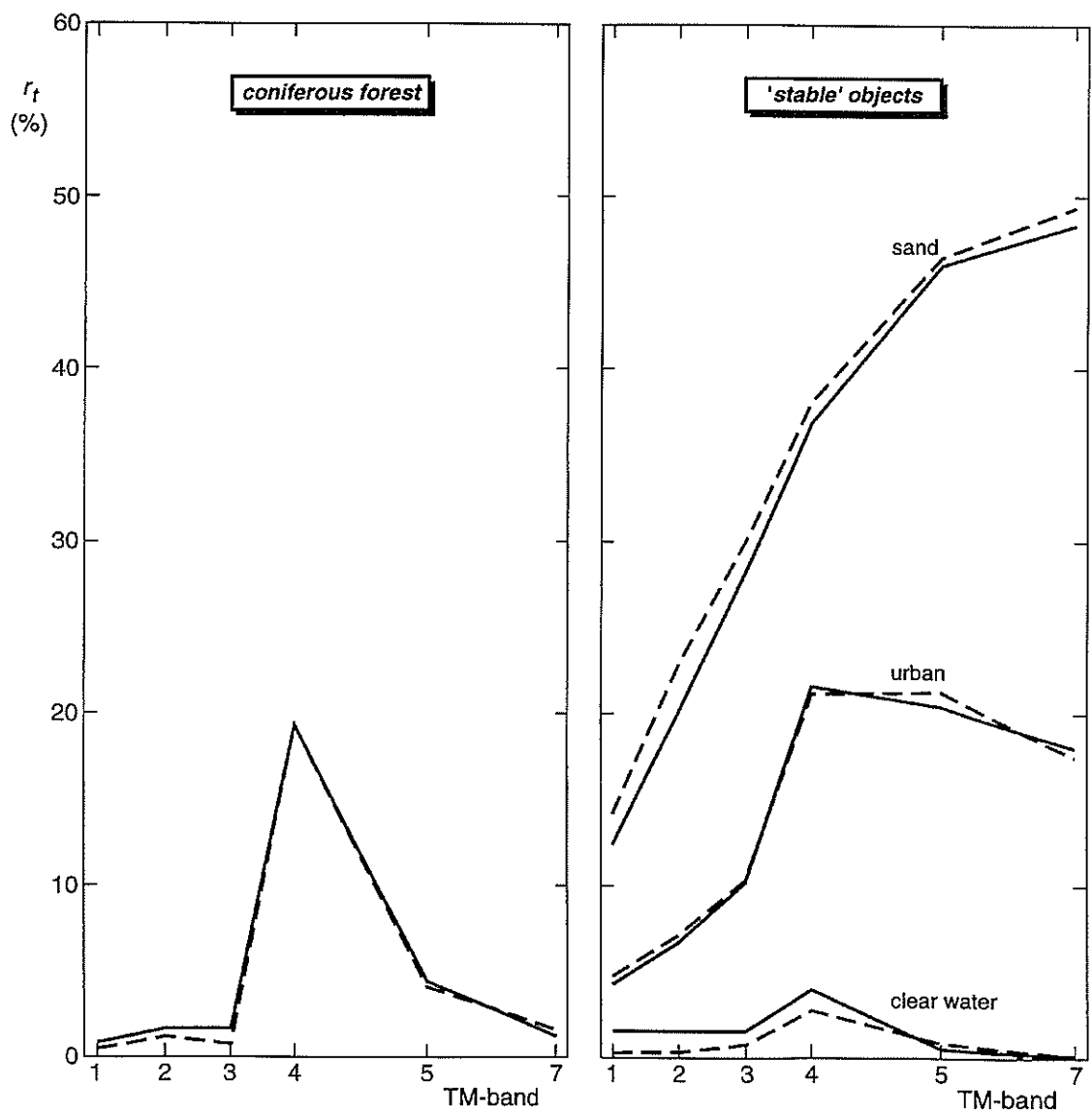


Fig. 9.9 TM-derived spectral reflectance for coniferous forest and other "stable" objects on the two dates



can be in any stage of development during the summer. For sugar beet one can safely assume that the growth stage on 3 August 1986 is similar to the one on 28 August 1973 so that the TM-derived spectrum of 3 August should be comparable with the NIWARS measurement. Especially for TM5 and TM7 correspondence is very good, so there is no evidence of water vapour absorption in these bands, like was found by Epema, as otherwise the reflectances in TM5 and TM7 would be significantly smaller than the measured values, and this is not the case. There is, however, a remarkable difference in TM2: the measured spectrum shows a pronounced peak in the reflectance in the green, whereas this peak in the TM-derived spectrum is much weaker. Two possible causes of this are 1) the band width of 85 nm associated with TM2, which could be too wide to resolve the green peak clearly, and 2) the adjacency effect, which tends to wash out spectral differences between a target and its surroundings. More evidence for the latter is found in Fig. 9.9, which shows TM-derived spectra of coniferous forest and other 'stable' objects like sand, an urban area and a small lake containing clear water. This evidence appears in the form of a peak in the reflectance in TM4 for the small lake and the fact that the spectra of sand and clear water are wider apart on 3 August than on 16 June, which can be explained by a stronger adjacency effect on 16 June due to more haze in the atmosphere. Nevertheless, the performance of the atmospheric correction in terms of spectral characterization can be considered good, as the changes found for the stable objects between the two dates are relatively small. Without atmospheric correction the difference between the two dates is greatest for dark objects, especially in the blue band (TM1), as can be seen by comparison of the darkest pixel values in both Tables. For an urban area and coniferous forest the mean differences in *DN*-values between the two dates are listed below. This also gives substantial differences.

Object	Date	TM1	TM2	TM3	TM4	TM5	TM7
Urban	16-06	101.2	43.2	43.8	65.7	72.1	36.5
	03-08	83.3	35.9	37.4	57.4	63.6	31.7
Forest	16-06	82.7	31.8	24.7	69.8	30.7	9.5
	03-08	62.0	23.0	17.6	59.8	26.2	8.2



Table 9.1 Darkest pixel correction results for 16-06-1986

Sun-earth distance in AU : 1.01593661

Solar zenith angle (deg) : 33.7

TM	λ (nm)	Darkest pixel	r_p	r_t	b_R	b_{O3}	b_A^*	b_A
1	485	78	.115	.000	.165	.008	.745	.743
2	560	28	.083	.010	.092	.030	.681	.675
3	660	22	.060	.008	.047	.010	.619	.604
4	830	11	.033	.000	.019	.000	.518	.518

α = -.671

β = .458

R^2 = .995

RMSE = .007

Correction constants:

TM	ρ_{so}	$T_1 T_2$	ρ_{dd}
1	.1150	.7188	.2025
2	.0750	.7567	.1451
3	.0524	.8479	.1017
4	.0333	.9136	.0670
5	.0115	.9646	.0287
7	.0085	.9718	.0232



Table 9.2 Darkest pixel correction results for 03-08-1986

Sun-earth distance in AU : 1.01470556

Solar zenith angle (deg) : 39.6

TM	λ (nm)	Darkest pixel	r_p	r_t	b_R	b_{O3}	b_A^*	b_A
1	485	59	.093	.000	.165	.008	.457	.457
2	560	20	.063	.007	.092	.030	.406	.401
3	660	15	.043	.007	.047	.010	.348	.345
4	830	7	.020	.000	.019	.000	.218	.208

α = -.911

β = .236

R^2 = .999

RMSE = .003

Correction constants:

TM	ρ_{so}	$T_1 T_2$	ρ_{dd}
1	.0933	.7519	.1800
2	.0566	.7872	.1231
3	.0363	.8797	.0782
4	.0202	.9432	.0452
5	.0051	.9831	.0140
7	.0033	.9879	.0101



9.8 Conclusions

A four-stream atmospheric radiative transfer model has been described which has been applied to the correction of Landsat Thematic Mapper images for atmospheric effects and the calibration in surface reflectance units.

The proposed method makes use of literature data and the darkest pixels in a scene in order to derive the aerosol optical thickness as the prime unknown quantity. This has the advantage that it is easy to implement the algorithm on an operational bases, since the necessary input is extracted from the image and further consists of usually available data like the calendar date and the solar elevation at the time of overpass. A disadvantage is that violation of the assumptions made leads to errors.

It has been demonstrated that application of this simple method of atmospheric correction gives satisfactory results in most cases. Samples of the surface reflectance extracted from the corrected images appear to correspond rather well with the results of reflectance measurements in the field. However, there are indications that in the infrared the absorption by water vapour in the atmosphere is sometimes underestimated. Moreover, for accurate results it might be worth considering a more advanced modelling of the adjacency-effect.



10 A COUPLED SOIL-VEGETATION-ATMOSPHERE MODEL

10.1 Introduction

The observation of vegetation canopies by means of optical remote sensing equipment on board airborne or spaceborne platforms can be simulated by a combined model which has been constructed from elements discussed in previous chapters. Using such a combined model has the advantage that the effects of a large number of parameters on the radiance measured at the sensor can be studied, and this may help explain many of the effects observed in images of vegetated areas. The combined model is entirely based on four-stream radiative transfer theory, and although this may not be sufficiently accurate for the atmospheric effects, the degree of realism in the model is still considered high enough for the atmospheric conditions under which optical remote sensing can successfully be applied anyway. Effects included in the model are, for instance, the bidirectional nature of surface reflectance (for vegetation), the adjacency effect and the bidirectional scattering behaviour of aerosols and air molecules. The flying altitude of the platform is accounted for by dividing the atmosphere in a layer below the aircraft and one above it.

The physical quantity calculated by the model is the so-called normalized spectral radiance ν , defined by $\nu = \pi L_o / E_s^o$, where E_s^o is the extraterrestrial solar spectral irradiance on a plane perpendicular to the sunrays and L_o is the spectral radiance at the sensor in the viewing direction.

Observation of a white Lambertian planet without an atmosphere at the location where the sun is at the zenith would result in a normalized spectral radiance of one. In this chapter the model, which is called OSCAR (from Optical Soil-Canopy-Atmosphere Radiation), is discussed in more detail and some of its possible applications are illustrated.

10.2 Model design

Basic building blocks of the combined model are the SAILH subroutine for calculation of the reflectances and transmittances of a single vegetation canopy layer, with incorporation of the hot spot effect, and a similar subroutine called SAM (Simple Atmosphere Model) for calculation of the



same quantities for an atmospheric layer. The earth's surface is supposed to exist of a uniform background and a small target. Both background and target are described by means of a soil reflectance, which is assumed to be Lambertian, and by vegetation canopy parameters. Soil and vegetation for background and target are allowed to be different from one another. The reflectances of the soil for target and background are called r_{st} and r_{sb} , respectively.

The four-stream radiation interactions between and within the elements atmosphere, background and target are illustrated by means of the flux interaction diagram of Fig. 10.1. In this diagram each square represents an incident flux and each circle on exitent flux. The solid arrows indicate the direction of flow and the reflectances (r, ρ) or transmittances (τ) are the multiplication factors to be applied. Interfaces between layers or those between layers and surfaces are connected by dashed arrows. For these the associated multiplication factor equals one.

Fig. 10.1 can be used to calculate $E_o = \pi L_o$ at the top of the atmospheric layer if the incident fluxes E_s and E^- at this level are given. For this, the adding method is used, by starting at the soils, and calculating equivalent reflectances for the top each time a vegetation or atmospheric layer is added. However, when observation takes place inside the atmosphere, such as for remote sensing from airborne platforms, Fig. 10.1 only applies up to the atmospheric layer under the sensor. Therefore, it is necessary to split the atmosphere in a lower part under the sensor and an upper part which is above the sensor. For this, use is made of the height profiles also mentioned in chapter 9, from Sturm (1981). From these profiles the optical thickness for Rayleigh and aerosol scattering of both layers can be computed. As in the combined model also absorption by water vapour is included, for which only the total optical thickness is supposed to be given, it is necessary to assume an extinction profile for this component as well. Since water vapour, like aerosol, is concentrated at low altitudes, the aerosol scale height H_1 (for the layer below 5.5 km) is applied for water vapour too.

Absorption by ozone is supposed to take place in a layer at 25 km altitude, separate from the other atmospheric layers.

For the calculation of the normalized radiance at the level of the sensor, first the diagram of Fig. 10.1 is used, with atmospheric parameters of the lower atmosphere. This leads to four equivalent reflectances which apply to the combination background-target-lower atmosphere. These reflectances are:

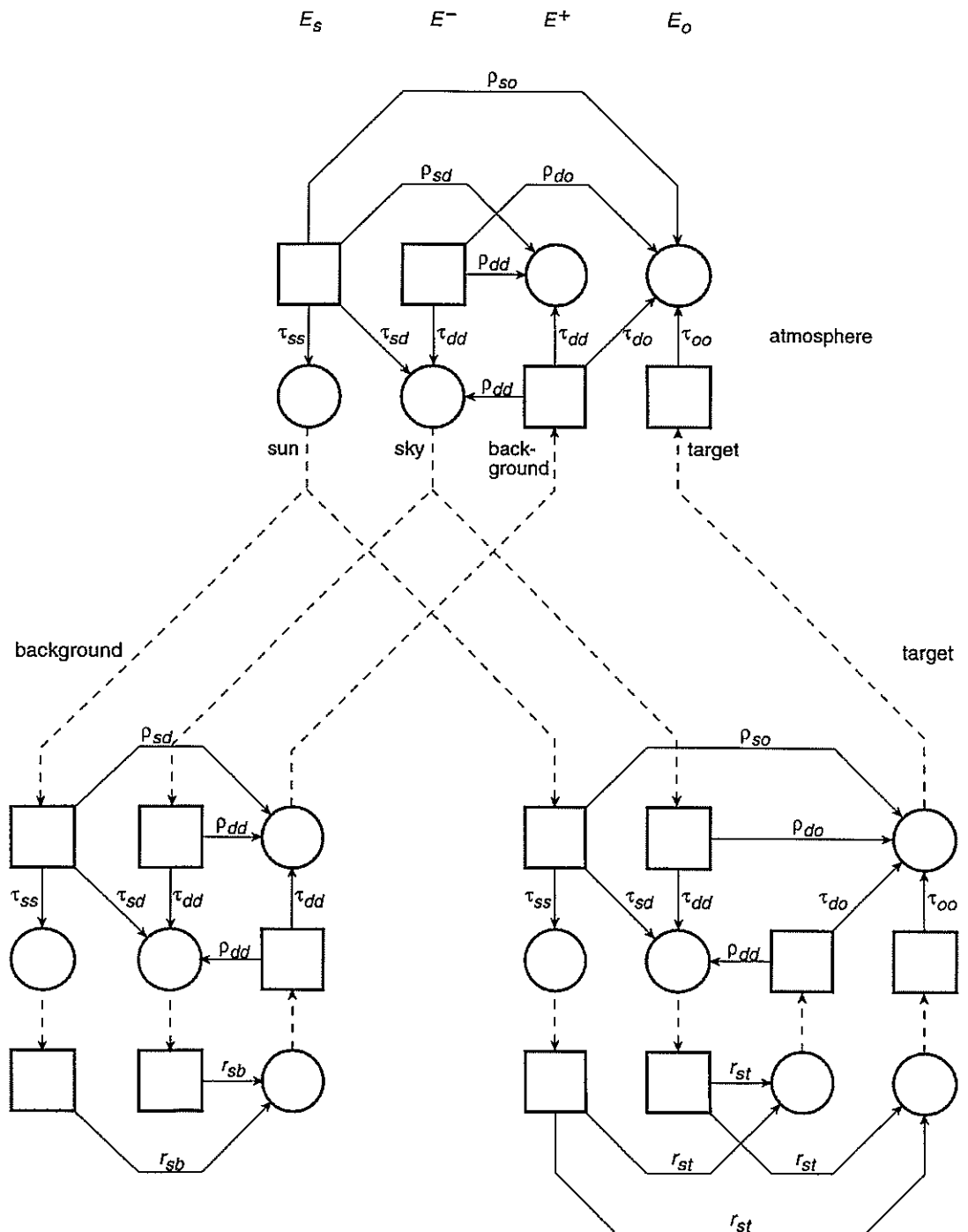


Fig. 10.1 Four-stream flux interaction diagram representing the observation of a non-Lambertian vegetation target through the atmosphere



- r_{ddb} = diffuse "background" reflectance for diffuse incidence,
- r_{sdb} = diffuse "background" reflectance for specular incidence,
- r_{dot} = directional "target" reflectance for diffuse incidence,
- r_{sot} = bidirectional "target" reflectance.

Quotes are used in order to indicate that the reflectances have been modified by the lower atmosphere. In Fig. 10.2 these reflectances are placed in a flux interaction diagram showing the interactions with the upper atmospheric layers.

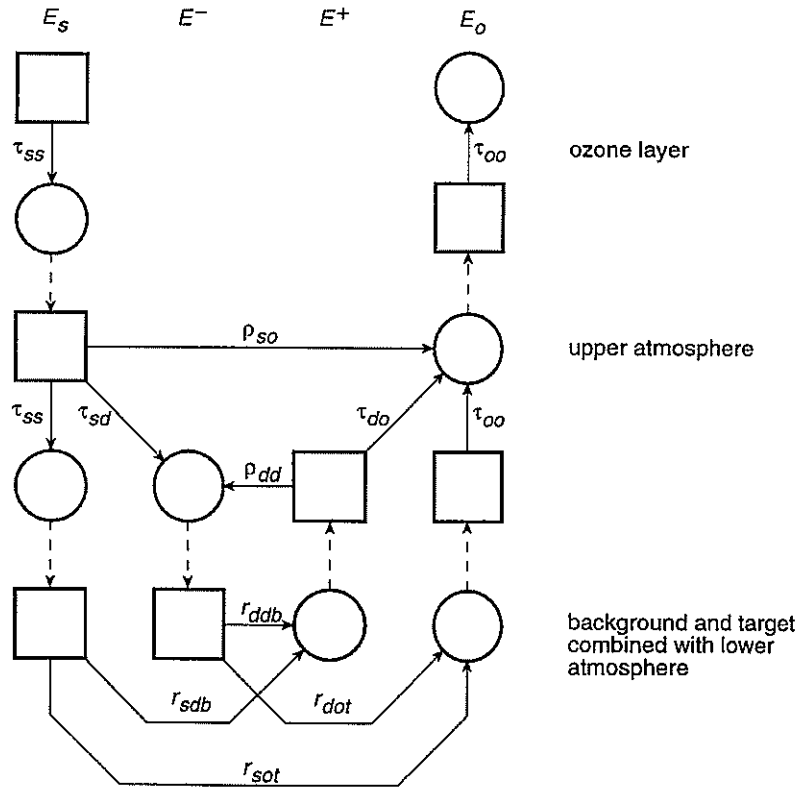


Fig. 10.2 Flux interaction diagram with split atmosphere. Background and target have been combined with the lower part of the atmosphere

From this diagram it becomes clear that the radiance at the top of the lower atmosphere can be found from

$$\pi L_o = E_o = r_{sot} E_{sun} + r_{dot} E_{sky}$$

where all quantities refer to the level at the top of the lower atmosphere. Here $E_{sun} = E_s^o \cos \theta_s \tau_{ssO_3} \tau_{ssu}$, where the subscripts O_3 and u refer to the



ozone layer and the upper atmospheric layer, respectively. E_{sky} is given by

$$E_{sky} = E_s^o \cos \theta_s \tau_{ssO_3} \frac{\tau_{sdu} + \tau_{ssu} r_{sdb} \rho_{ddu}}{1 - r_{ddb} \rho_{ddu}} .$$

For the normalized radiance $\nu = \pi L_o / E_s^o$ this gives

$$\nu = \cos \theta_s \tau_{ssO_3} \left[\tau_{ssu} r_{sot} + \frac{\tau_{sdu} + \tau_{ssu} r_{sdb} \rho_{ddu}}{1 - r_{ddb} \rho_{ddu}} r_{dot} \right] .$$

Except for the cosine-effect of the solar zenith angle, this expression contains only reflectances and transmittances.

For satellite observations the sensor is above the ozone layer, and in that case also the transmittance through this layer in the direction of view must be taken into account. As in that case there is no upper atmospheric layer to consider, the expression for the normalized radiance becomes

$$\nu = \cos \theta_s \tau_{ssO_3} r_{sot} \tau_{ooO_3} ,$$

where τ_{ooO_3} is the upward direct transmittance through the ozone layer, and r_{sot} is the bidirectional reflectance of the target, with incorporation of the complete atmosphere.

Model output results can easily be compared with calibrated radiance measurements of airborne or spaceborne sensors. For this, it is only necessary to convert the normalized radiances into actual radiances by

$L_o = \nu E_s^o / \pi$, where E_s^o in this case is the extraterrestrial spectral solar irradiance in $W/(m^2 \mu m)$ in the associated spectral band of the sensor.

10.3 Model parameters

The parameters of the OSCAR model can be divided into groups for the categories background, target, atmosphere and observational as follows:

Background parameters:

LIDF = leaf inclination distribution function

LAI = leaf area index



ρ_L = single leaf reflectance
 τ_L = single leaf transmittance
 ρ_s = soil reflectance

Target parameters:

LIDF = leaf inclination distribution function
 LAI = leaf area index
 HOT = hot spot size parameter
 ρ_L = single leaf reflectance
 τ_L = single leaf transmittance
 ρ_s = soil reflectance

The hot spot size parameter is the horizontal correlation length (see chapter 8) divided by the canopy height. For the background it does not need to be specified because the background reflectance parameters r_{sdb} and r_{ddb} are not affected by it.

Atmospheric parameters:

V = sea level horizontal visibility
 λ = wavelength of the radiation
 α = aerosol Ångström coefficient
 ω_A = aerosol single scattering albedo
 b_{H_2O} = water vapour optical thickness
 b_{O_3} = ozone optical thickness

The wavelength λ must be specified in this case because the optical thicknesses of Rayleigh and aerosol scattering, the aerosol backscattering efficiency η_A and the aerosol scattering phase function $p_A(\delta)$ are computed by means of interpolation based on the wavelength.

Observational parameters:

h = observation height of the sensor
 θ_s = solar zenith angle
 θ_o = viewing zenith angle
 ψ = relative azimuth angle



The observation height might also be considered an atmospheric parameter, since the division of the optical thicknesses for Rayleigh scattering, aerosol scattering and water vapour absorption over the lower and upper atmospheric layers depends on it.

Fixed parameters and relationships assumed in the model are the following:

$$\begin{aligned}
 H_0 &= 8.5115 \text{ km (scale height for Rayleigh scattering)} \\
 H_2 &= 3.748 \text{ km (scale height for aerosol scattering above } h = 18 \text{ km)} \\
 \beta_{A5}(550) &= 0.0030765 \text{ km}^{-1} \text{ (aerosol extinction coefficient at 550 nm for } 5.5 \text{ km} \leq h \leq 18 \text{ km)} \\
 \beta_{R0}(550) &= 0.0116 \text{ km}^{-1} \text{ (Rayleigh extinction coefficient at 550 nm at sea level)} \\
 \beta_{A0}(550) &= (\ln 50) / V - \beta_{R0}(550) , \\
 H_1 &= 5.5 / \ln [\beta_{A0}(550) / \beta_{A5}(550)] , \text{ (scale height aerosol below } h = 5.5 \text{ km)} \\
 \beta_{R0}(\lambda) &= \beta_{R0}(550) (\lambda / 550)^{-4.069} \\
 \beta_{A0}(\lambda) &= \beta_{A0}(550) (\lambda / 550)^\alpha \\
 \beta_{A5}(\lambda) &= \beta_{A5}(550) (\lambda / 550)^\alpha \\
 \beta_{Rh}(\lambda) &= \beta_{R0}(\lambda) \exp(-h/H_0) \\
 \beta_{Ah}(\lambda) &= \beta_{A0}(\lambda) \exp(-h/H_1) \quad (0 < h < 5.5) \\
 \beta_{Ah}(\lambda) &= \beta_{A5}(\lambda) \quad (5.5 < h < 18) \\
 \beta_{Ah}(\lambda) &= \beta_{A5}(\lambda) \exp[(18-h)/H_2] \quad (h > 18) \\
 \beta_{H2O}(0) &= b_{H2O}/H_1 \\
 \beta_{H2O}(h) &= \beta_{H2O}(0) \exp(-h/H_1)
 \end{aligned}$$

The above model of optical densities in the atmosphere is similar to the one described in chapter 9, except for the last two equations, which refer to water vapour. The constants and the fixed relationships make it possible to estimate the optical thicknesses for Rayleigh scattering, aerosol scattering and water vapour absorption for any layer in the atmosphere, when only the parameters V , λ , α and b_{H2O} are given.

For the leaf inclination distribution function use is made of the discrete frequency distributions of chapter 7, numbered 1 to 9, plus the discrete approximation of the spherical distribution, which is numbered 10. Each



distribution is characterized by leaf inclination frequencies at thirteen discrete angles (Verhoef, 1984).

For the aerosol phase function, reference is made to a phase function file, which contains the phase function at 34 scattering angles for five wavelengths.

As mentioned in chapter 9, cubic spline interpolation is applied to calculate the phase function at any scattering angle, and linear interpolation according to wavelength is used to estimate it for arbitrary wavelengths.

10.4 Implementation aspects

The model OSCAR has been coded as a FORTRAN subroutine with 21 input parameters and one output parameter (the normalized radiance). For the control of this large number of parameters a program was written in which the number of input parameters has been reduced to fifteen by introducing "types" for certain variables. In this way the leaf type, for instance, may indicate the spectral reflectance and transmittance data at all the wavelengths considered. The variables defined in this way are "soil", "leaf", "background" and "H2O". Here "soil" refers to the soil spectral reflectance at all wavelengths of interest, "leaf" to single leaf spectral reflectance and transmittance, "background" to soil type, leaf type, LAI and LIDF of the background, and "H2O" to the water vapour optical thickness at all wavelengths. A variable called "band" contains the wavelengths to be considered.

For the control of the input parameters of the model a variable looping mechanism has been designed. This means that the looping order and the number of states each variable is in are specified by means of an input file, making program modifications each time a new set of simulations is done unnecessary. For instance, simulations for varying viewing angle at several values of the solar zenith angle, and simulations for varying target LAI at several wavelengths are carried out by means of the same program, just by specifying the proper input file. In principle, the program carries out a nested loop on the input parameters of fifteen levels deep with arbitrary looping order. At the deepest level the OSCAR model is called and the calculated normalized radiance is collected for output.



10.5 Some examples of simulation results

In this section two aspects of the coupled model are investigated, namely the effect of the height of the platform of observation and the adjacency effect. Before presenting the results for realistic vegetation targets and backgrounds, first the functioning of the model for some extreme cases, namely black (zero reflectance) and white Lambertian (100 % reflectance) bare soils, is demonstrated for a varying observation height in the atmosphere.

Fig. 10.3 shows the vertically upward normalized radiance ν at 550 nm as a function of the height above sea level, h , for the four combinations of black and white targets and backgrounds, for a) $V = 40$ km and b) $V = 5$ km, as calculated by means of the OSCAR model. The solar zenith angle θ_s is 45° . The aerosol single scattering albedo ω_A is one in this case. The meaning of the symbols in Fig. 10.3 is as follows:

open triangles	:	white target, white background
open circles	:	white target, black background
solid triangles	:	black target, white background
solid circles	:	black target, black background

From this illustration it appears that even for a clear atmosphere ($V = 40$ km) the adjacency effect, as expressed by the difference between the circles and the triangles, is already considerable. For a hazy atmosphere (Fig. 10.3.b) the effect above some level (2 km) in the atmosphere is so strong that a black target with a white background gives a higher radiance than a white target with a black background.

Fig. 10.4 is similar to Fig. 10.3, except for the aerosol single scattering albedo, which is 0.75 in this case. The effect of this is substantial, especially for $V = 5$ km. All radiance levels are lower due to the absorption of radiation by the atmospheric aerosol. While in Fig. 10.3 the combination white target - white background gives an almost constant upward radiance (independent of the observation height), in Fig. 10.4 the upward radiance decreases with height for this case. This can be explained by the fact that in this situation a white earth is observed through an increasingly thick layer of absorbing material, which leads to a decreasing apparent brightness of the earth's surface. As to the height dependence, in general it can be stated that this effect is concentrated in the lower 5 km of the atmosphere, with the strongest variation close to the surface. Airborne observations from 10 km

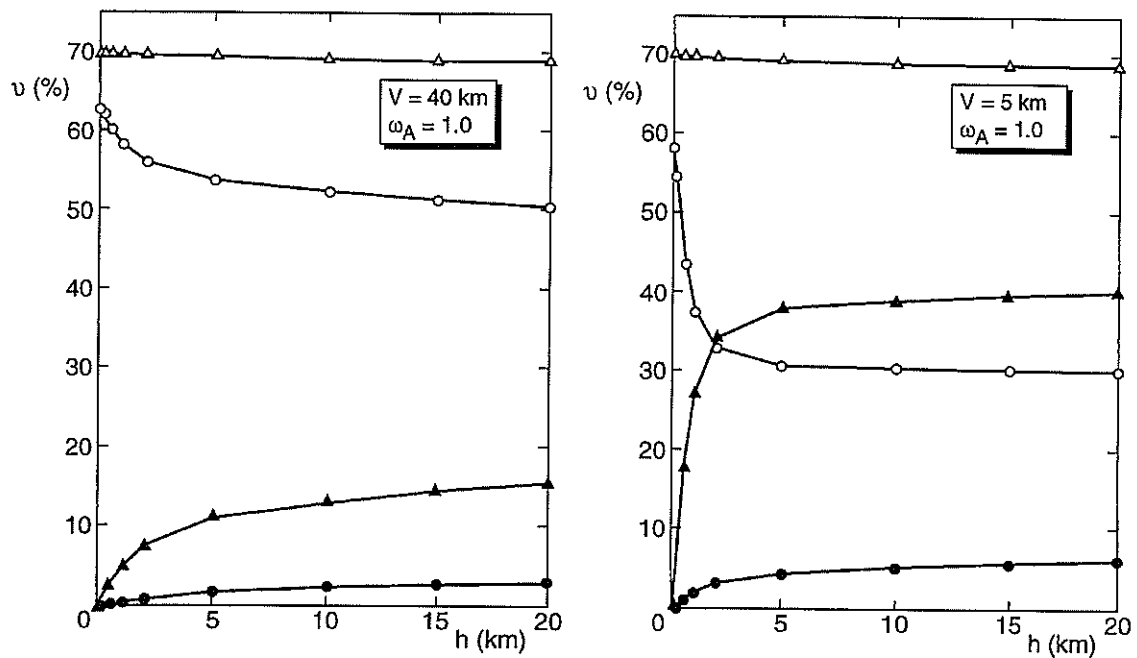


Fig. 10.3 Normalized vertical radiance at 550 nm as a function of the height for black and white target and background. $\omega_A = 1.0$; a) $V = 40$ km, b) $V = 5$ km

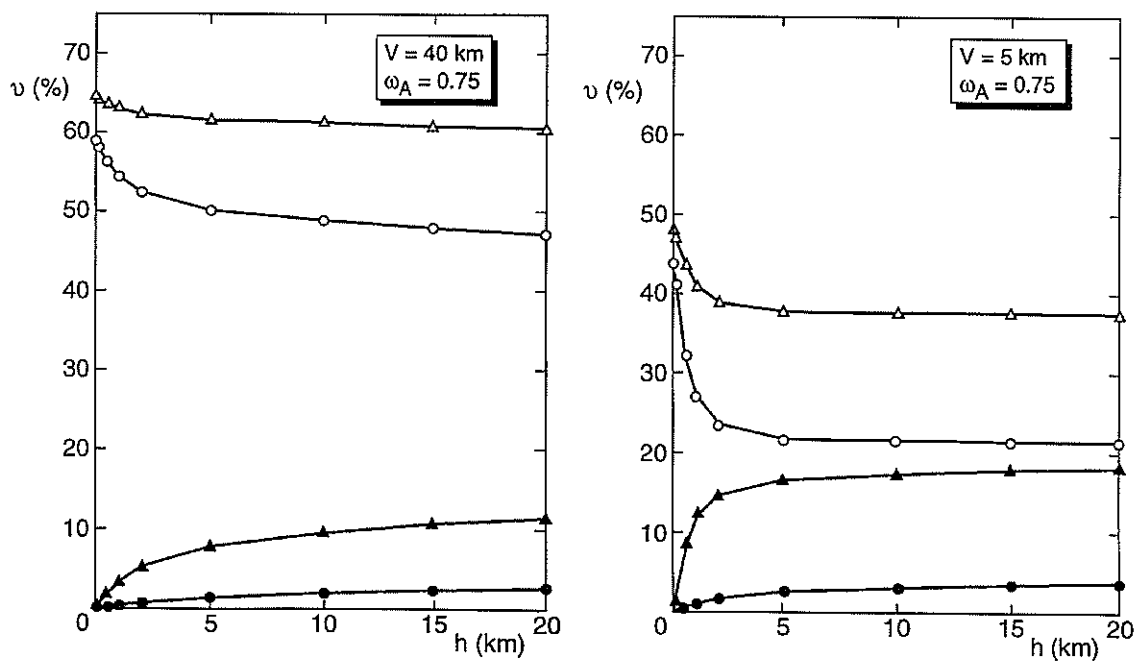


Fig. 10.4 As Fig. 10.3, but $\omega_A = 0.75$



altitude will already differ little from satellite observations.

In the remainder of this section directional normalized spectral radiance profiles in the principal plane ($\psi = 0^\circ, 180^\circ$) are shown which simulate the observation of vegetation canopies in the red at 670 nm and in the near infrared at 830 nm under different atmospheric conditions. Ground based measurements from 10 m height and satellite observations from 800 km altitude are simulated for hazy ($V = 5$ km) and clear ($V = 40$ km) atmospheric conditions. The background consists either of bare soil or of dense vegetation with an LAI of 8 and a plagiophile leaf angle distribution (LIDF no. 4 of chapter 7). The constant parameters are given by:

$$\begin{aligned}\theta_s &= 30^\circ ; \\ \omega_A &= 0.95 ; \\ \text{HOT} &= 0.1 .\end{aligned}$$

Spectral parameters:

Parameter	$\lambda = 670$ nm	$\lambda = 830$ nm
ρ_s	0.175	0.280
ρ_L	0.075	0.520
τ_L	0.007	0.440
$b_{\text{H}_2\text{O}}$	0.00	0.05
b_{O_3}	0.01	0.00

The next Table shows the variable parameters pertaining to the associate Figure numbers. The entry "var" refers to a series of values of the parameter.



Figure	Backgr.	LAI	LIDF	V (km)	<i>h</i>	λ (nm)
10.5	bare soil	var	4	5	10 m	670
10.6	bare soil	var	4	40	10 m	670
10.7	bare soil	var	4	40	10 m	830
10.8	bare soil	0.5	var	40	10 m	670
10.9	bare soil	4	var	40	10 m	670
10.10	bare soil	var	4	40	800 km	670
10.11	dense veg.	var	4	40	800 km	670
10.12	bare soil	var	4	40	800 km	830
10.13	dense veg.	var	4	40	800 km	830

Fig. 10.5 shows the directional profiles in the red at 670 nm as a function of the zenith observation angle in the principal plane for vegetation canopies with LIDF no. 4 and LAI values of zero, 0.25, 0.5, 1, 2 and 4. The background is bare soil and the visibility $V = 5$ km (hazy). The observation height is 10 m above the ground. At this height there is already a small atmospheric effect for large zenith viewing angles, as evidenced by the increase of the radiance for $LAI = 0$ (bare soil). As the soil is assumed to be Lambertian, this increase can only be attributed to the atmospheric effect. The curves for other values of the LAI all clearly show the hot spot effect. Figure 10.6 is similar to Fig. 10.5, except for the visibility, which is 40 km for the former. Comparing both, one sees that for the clear atmosphere the hot spot effect is substantially stronger than for a hazy atmosphere. In this Figure also the curve for $LAI = 8$ has been included in order to show the effect of an increasing radiance when the LAI increases from 4 to 8.

In Fig. 10.7 the conditions are the same as in Fig. 10.6, but the normalized radiance values are for the near infrared at 830 nm in this case. Unlike the situation in the red, in the near infrared a higher LAI always gives a higher radiance at all observation angles, and also the differentiation of LAI at high values of this parameter is much better. The above simulation results have much in common with those shown in chapter 8, which showed bidirectional

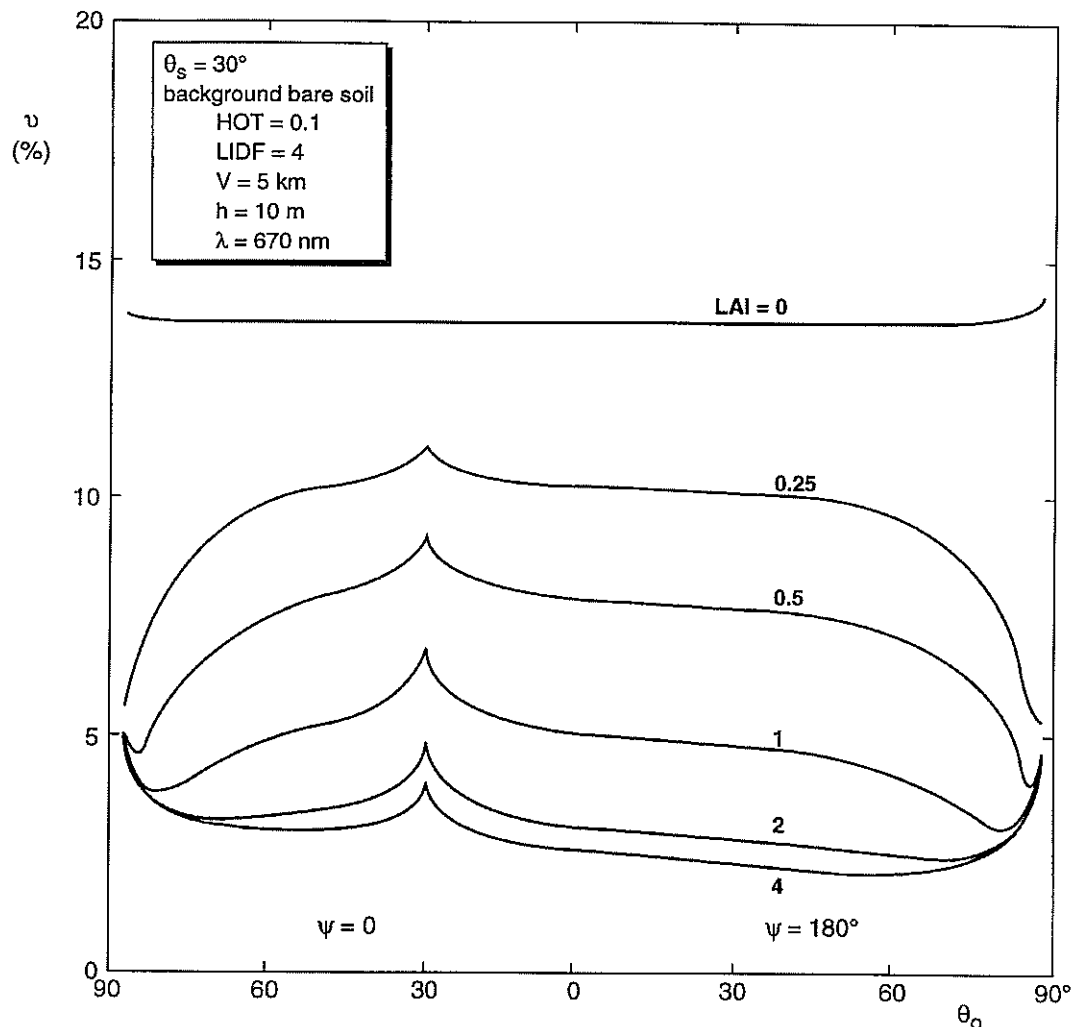


Fig. 10.5 Normalized radiance profiles for $h = 10$ m at $\lambda = 670$ nm and $V = 5$ km. Bare soil background

reflectances instead of normalized radiances. Note however, that in chapter 8 it was necessary to specify the ratio of diffuse to total irradiance for each wavelength, whereas with the OSCAR model the effect of the diffuse downward irradiance is automatically accounted for by means of the atmospheric part of the model.

Figures 10.8 and 10.9 show profiles of the normalized radiance in the red for varying LIDF at LAIs of 0.5 and 4, respectively. At small LAI (Fig. 10.8) discrimination of the LIDF is best at nadir viewing angle. Minimum sensitivity to the LIDF occurs at about 60° zenith angle, both in the forward scattering part and the backscattering part of the principal plane. Also in the

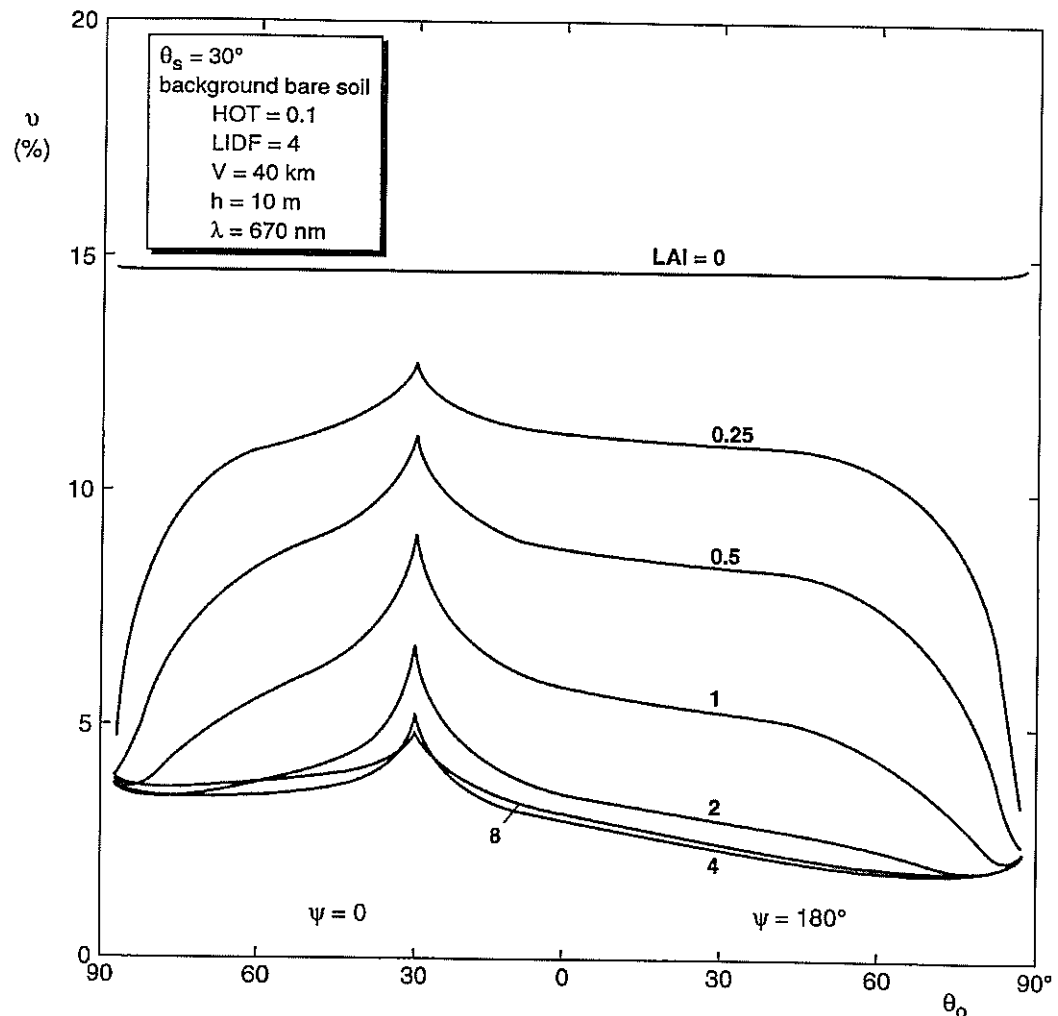


Fig. 10.6 Normalized radiance profiles for $h = 10$ m at $\lambda = 670$ nm and $V = 40$ km. Bare soil background

hot spot a minimum sensitivity to the LIDF is found.

For an LAI of 4 (Fig. 10.9) the situation is quite different, however. In this case the sensitivity to the LIDF is especially high in the forward scattering part of the principal plane ($\psi = 180^\circ$), and relatively low in the backscattering part, in particular at the hot spot point. Both examples demonstrate the complexity of the relationships between the morphological canopy parameters LAI and LIDF on one hand and the observed directional radiances in the red and the near infrared parts of the spectrum on the other.

For $\lambda = 670$ nm (red light), Figures 10.10 and 10.11 show directional normalized radiance patterns at satellite height under clear atmospheric

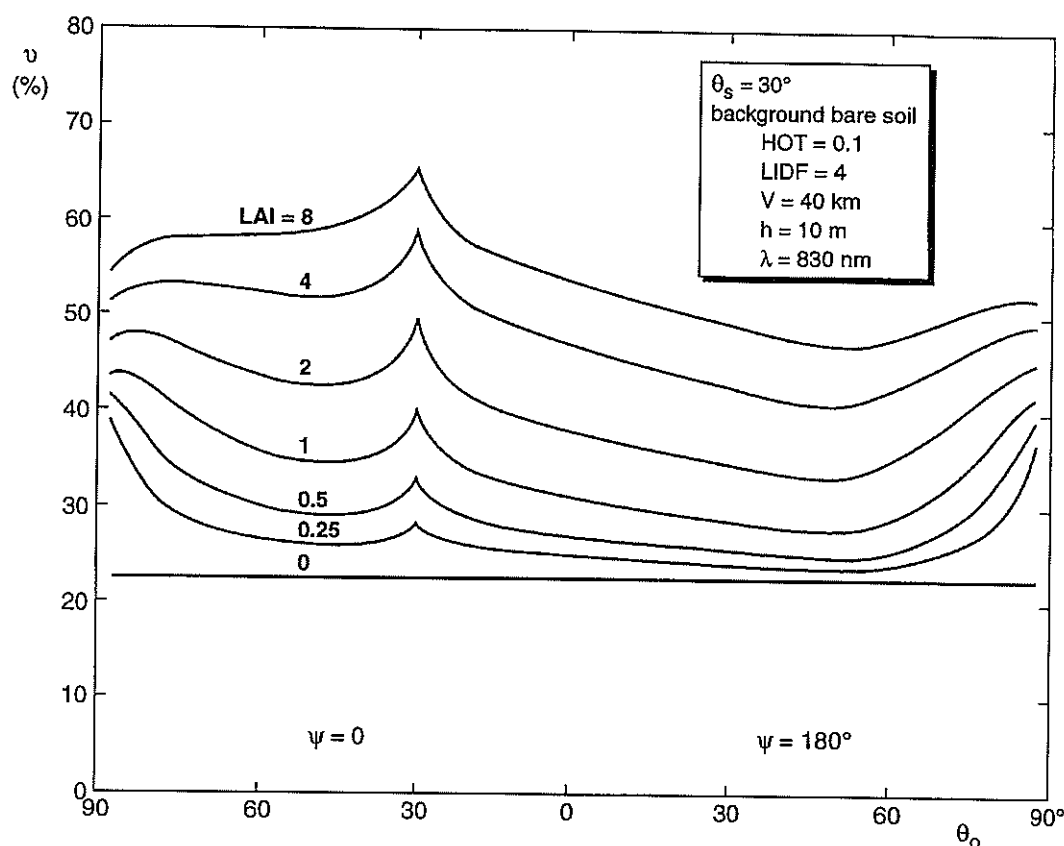


Fig. 10.7 Normalized radiance profiles for $h = 10$ m at $\lambda = 830$ nm and $V = 40$ km. Bare soil background

conditions. In Fig. 10.10 the background is bare soil. In this case the surroundings are much brighter than the observed target, and this makes that the radiances increase considerably for large zenith observation angles, when the adjacency effect is stronger. This is especially the case for $LAI = 4$, since the contrast between target and background is maximum then. Both Figures demonstrate that the hot spot effect remains clearly visible through the atmosphere for small (0.25) as well as large LAI. At $LAI = 0$ (bare soil) there is no canopy hot spot effect, but the aerosol scattering in the glory region produces a ripple around the direction of exact backscattering. From this it becomes clear that both effects (aerosol glory and canopy hot spot) are mixed in the case of spaceborne observations.

In Fig. 10.11 the only difference with Fig. 10.10 is the background, which is dense vegetation now. Comparing both, one sees that the background is able to modify the directional radiance patterns considerably, both in shape and level. In Fig. 10.11 the background is dark in the red part of the



spectrum, and therefore the strong increase of the radiance at larger zenith angles does not occur. Also the radiance levels in general are lower because the surroundings of the target contribute less in this case.

Similarly to Figures 10.10 and 10.11, the effects in the near infrared are shown in Figures 10.12 (bare soil background) and 10.13 (dense vegetation background). Here the differences due to variation of the background seem to be less dramatic than in the red, but at large zenith angles the effect is again substantial. However, also for moderate zenith angles there is a significant difference in level between both simulations.

These results demonstrate that, even under clear atmospheric conditions, it will be difficult to reliably estimate the LAI from satellite radiance observations in the red and the near infrared without taking the surroundings of the target into account.

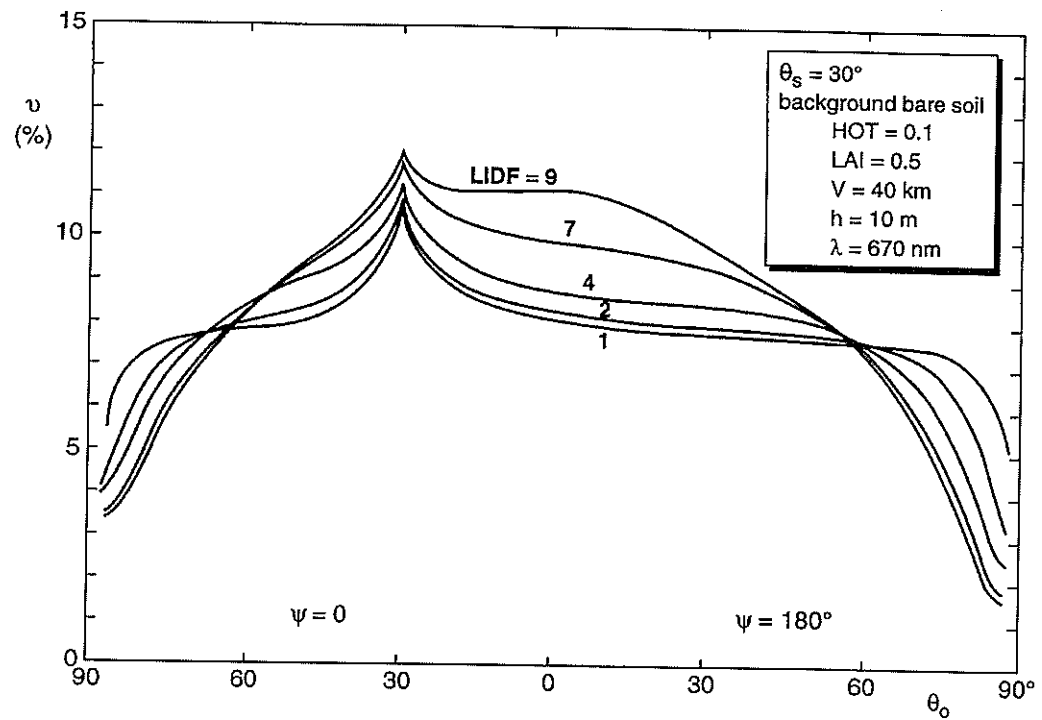


Fig. 10.8 Normalized radiance profiles for $h = 10$ m at $\lambda = 670$ nm and $V = 40$ km for different LIDFs at $LAI = 0.5$. Bare soil background

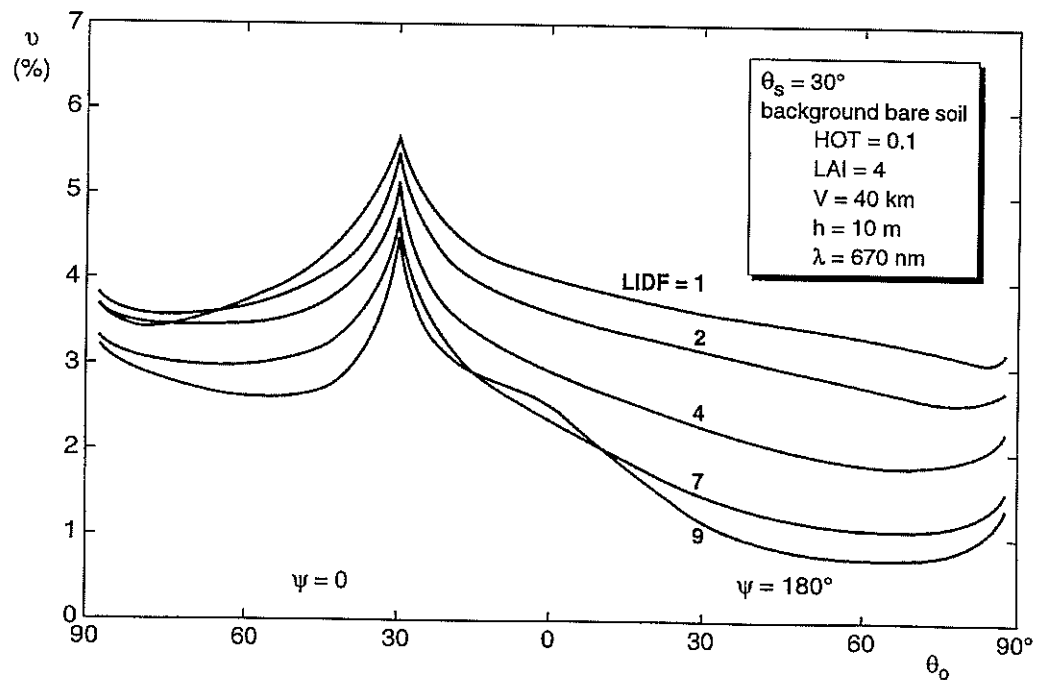


Fig. 10.9 Normalized radiance profiles for $h = 10$ m at $\lambda = 670$ nm and $V = 40$ km for different LIDFs at $LAI = 4$. Bare soil background

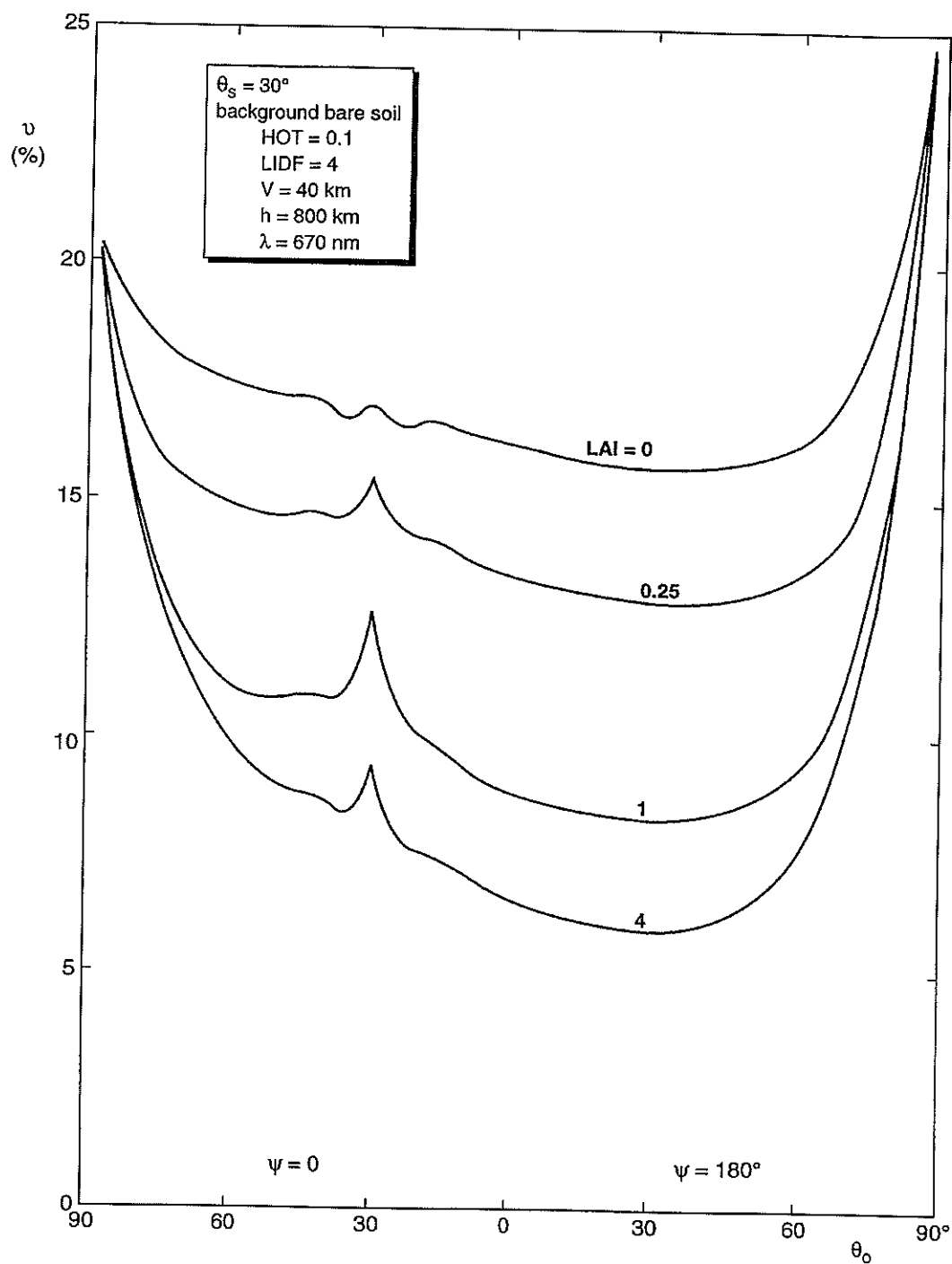


Fig. 10.10 Normalized radiance profiles at satellite height for $\lambda = 670$ nm and $V = 40$ km. Bare soil background

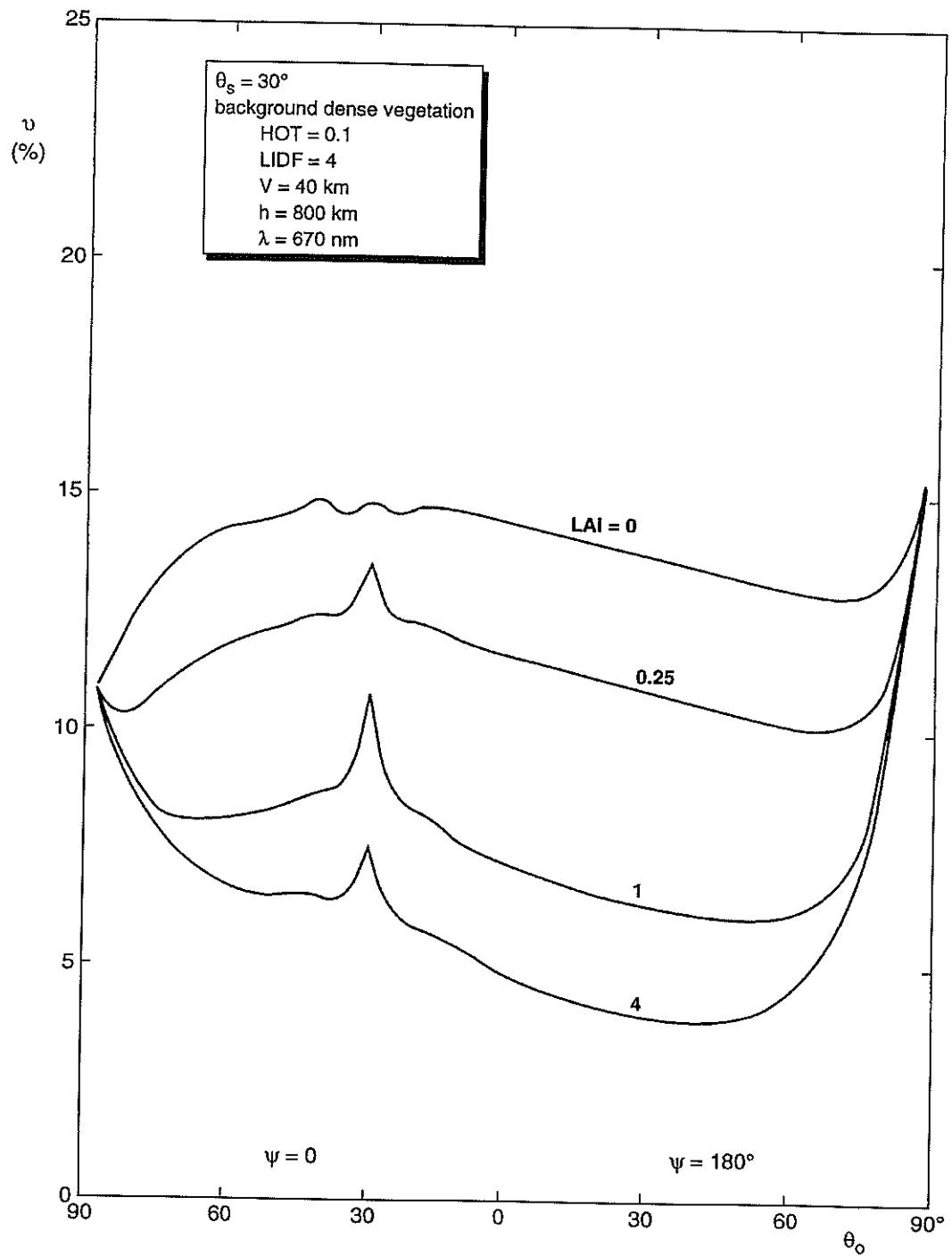


Fig. 10.11 Normalized radiance profiles at satellite height for $\lambda = 670 \text{ nm}$ and $V = 40 \text{ km}$. Dense vegetation background

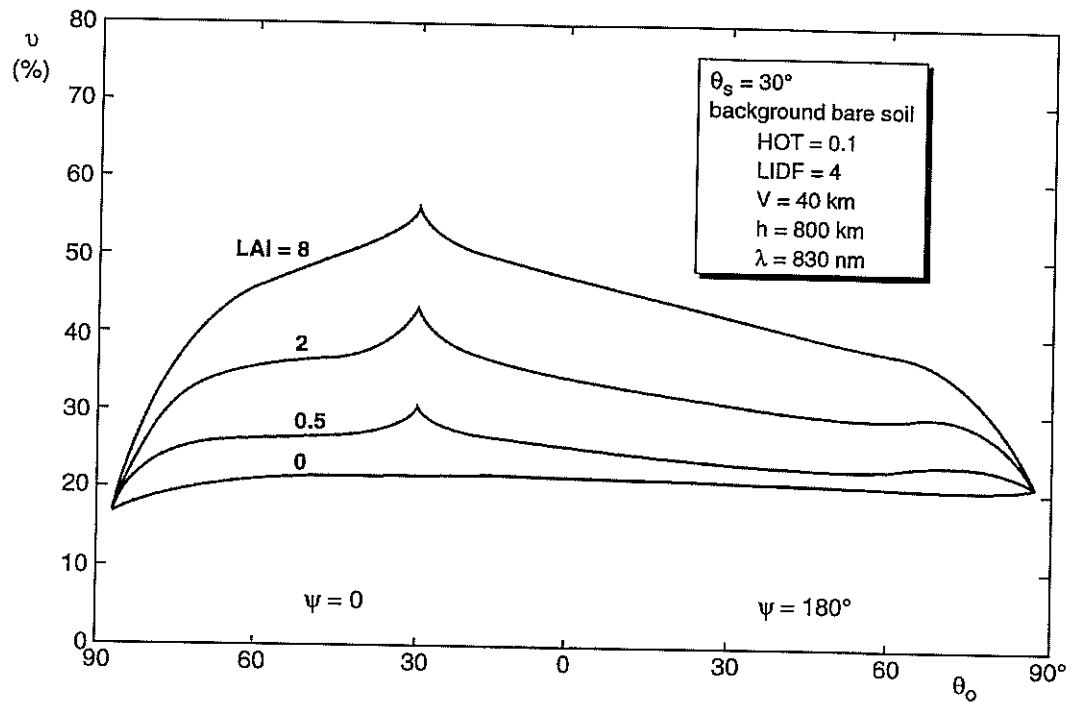


Fig. 10.12 Normalized radiance profiles at satellite height for $\lambda = 830 \text{ nm}$ and $V = 40 \text{ km}$. Bare soil background

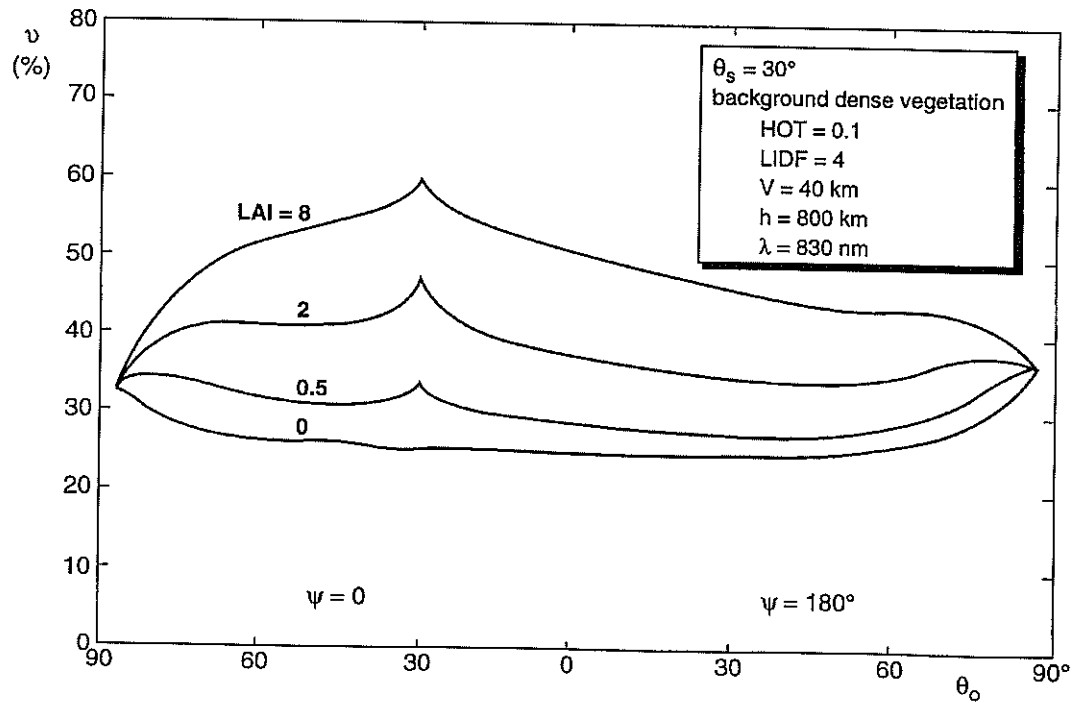


Fig. 10.13 Normalized radiance profiles at satellite height at $\lambda = 830 \text{ nm}$ and $V = 40 \text{ km}$. Dense vegetation background



10.6 Fish-eye views

By means of the OSCAR model a number of so-called fish-eye views of the directional patterns of normalized radiances have been produced in order to illustrate these for different solar zenith angles, spectral bands and observation heights. Fish-eye views are intended to show these variations as a function of all viewing directions in a hemisphere by means of circular discs in which the distance from the centre represents the zenith viewing angle and the radial to the centre the azimuth. For the cover of this thesis, a number of these discs have been generated and recorded on film as colour composites by means of the Optronics digital film recorder of NLR. All data were generated for a clear atmosphere with 40 km visibility.

The front cover shows the radiance pattern at satellite height in the blue, green and red parts of the spectrum in order to obtain an approximate true colour rendition. The sun is supposed to be positioned on the left at an elevation of 45° . In that case the combination of the canopy hot spot effect and the atmospheric glory region appears on the right as a bright spot surrounded by a slightly darker ring. The canopy LAI is one and the LIDF is spherical, for target as well as background. The colours observed are associated with variations of the amount of visual vegetation cover and the increasing atmospheric effects at larger viewing angles. The outer edge of the disc represents the horizon and the colour here is determined mainly by the absorption spectrum of the ozone layer, which is maximum in the green, so that the complementary colour (purple) is observed.

The back cover shows the sky radiance patterns on the left, the upward radiances observed just above the ground in the middle, and the upward radiances above the atmosphere on the right. The sky radiances are shown for the visible wavelengths in order to illustrate the blue colour of the sky. The upward radiances are shown in false colour (green, red and near infrared radiances represented in blue, green and red, respectively) in order to illustrate the usual colour rendition of satellite images in these spectral bands for e.g. the SPOT HRV instrument.

The solar zenith angle is 35° for the top row, 55° for the middle row and 75° for the bottom row. These angles are representative for satellite observations over the Netherlands at 11 hours local solar time in the summer, in spring or autumn, and in winter, respectively.

The sky radiance patterns are characterized by the very bright region around the sun caused by the aerosol forward scattering peak (aureole) and a



brightening towards the horizon. These regions are also least coloured. In the darker regions Rayleigh scattering dominates and the colour is blue due to the strong wavelength dependence of this type of scattering. For low solar elevation (lower left disc) the aureole becomes slightly yellow.

The directional radiances just above the ground clearly demonstrate the increasing proportion of vegetation (red colours) observed for more oblique viewing angles. For low solar elevation the centre of the disc becomes very dark because much shadow is observed in that case. The hot spot effect is strongest for high solar elevation and is accompanied by a change of colour because this effect is stronger in the visible than in the near infrared and because much sunlit soil is observed.

Above the atmosphere the colours observed on ground level have considerably bleached by the atmospheric effect. Also the hot spot effect has been reduced but it remains visible and gets mixed up with the aerosol glory region. However, the most striking feature in these discs is the very bright region on the left, which is caused by the strong forward scattering of the aerosol particles. Especially at low solar elevation this effect is very strong and the images of the associated discs were allowed to go into saturation for these regions on the left, in order to prevent the other parts from becoming too dark. Note that in the centre bottom image (ground level and 75° solar zenith angle) a beginning of this effect is already visible. This is caused by the fact that the actual height taken as input for the model was 10 m. So already at 10 m height above the ground an atmospheric effect is present when viewing is nearly horizontal in the direction of the sun.

11 APPLICATIONS

In this chapter it is shown how the application of radiative transfer modelling to remote sensing of vegetation can contribute to enhancement of the insight in certain problems and may lead to extraction of more useful information from remotely sensed images. One example of an application of radiative transfer modelling has been demonstrated already in chapter 9, namely correction of atmospheric effects. The applications discussed in this chapter are more related to the extraction of object information from optical remote sensing data.

Models are an abstraction of reality and always a compromise between realism and simplicity. Nevertheless, the use of models is attractive because it can partially replace the expensive and sometimes laborious execution of experiments, provided the model does represent reality sufficiently well. In the application of models three categories can be distinguished, namely prediction, explication and inversion.

Model predictions are carried out in order to explore sensitivities to input parameters, to find favourable conditions of measurement and to test certain data processing algorithms.

In the explication category a model simulation result is compared with the result of an actual experiment and it is shown that the model explains to some extent what was observed in reality. In this way the use of the model supports the interpretation of the experimental results and on the other hand this may lead to necessary improvements of the model when experimental and simulated results disagree to an unacceptable extent.

Model inversion is the estimation of model input parameters from experimental data by modifying the input of the model until its output sufficiently matches the measurements. This technique can be applied for instance to derive important biophysical canopy parameters like LAI from remotely sensed reflectance data.

In the following sections examples and possibilities of all three categories will be discussed.



11.1 NOAA satellite NDVI simulation

The most widely used vegetation index is the so-called normalized difference vegetation index (NDVI), defined by $NDVI = (I - R) / (I + R)$, where I is the radiance or reflectance in the near infrared and R the reflectance or radiance in the red. This vegetation index is a measure of vegetation development because it has been observed that during growth the near infrared radiance from crops increases while the red radiance decreases. As a result, the difference increases while the sum of both is more or less constant. The advantage of dividing the difference by the sum is that this gives to some extent compensation of effects due to varying solar zenith angle, soil brightness differences and leaf slope variation, since these effects have about an equal influence on difference and sum and thus cancel out. Maps of the NDVI on continental and global scale are produced on an operational basis from NOAA-AVHRR satellite data and have found application in the study of global vegetation dynamics and in monitoring the food situation in Africa.

The NDVI as determined from satellite radiance measurements is influenced by clouds, atmospheric haze and viewing angle effects. However, these effects all lead to an increase of mainly the red radiance, so that the influence on the NDVI is always in negative direction. The NOAA-AVHRR instrument enables daily observation of the same location on earth and the usual procedure for making NDVI maps is to apply the so-called maximum value compositing (MVC) technique, which consists of determining for each location the maximum NDVI over a period of one week, ten days, or a month. In this way most influences due to clouds, haze and oblique viewing are removed, since the maximum observed NDVI in a period is likely to represent the most favourable measurement condition.

To what extent the NDVI is a measure of vegetation development can be investigated by using a stand-alone canopy reflectance model. How the (top of atmosphere) NDVI is influenced by atmospheric haze, solar elevation and the viewing angle can be investigated by means of a coupled vegetation-atmosphere radiative transfer model.

An example of the latter is presented in Fig. 11.1. This plot shows the simulated NDVI - time profile for the year 1991 as it would be obtained from the NOAA-11 AVHRR data for an agricultural area in the Flevoland province in the Netherlands (Van Dijk *et al.*, 1994). In this case the crop growth model WDUET (Huygen, 1988) was used to generate daily values of

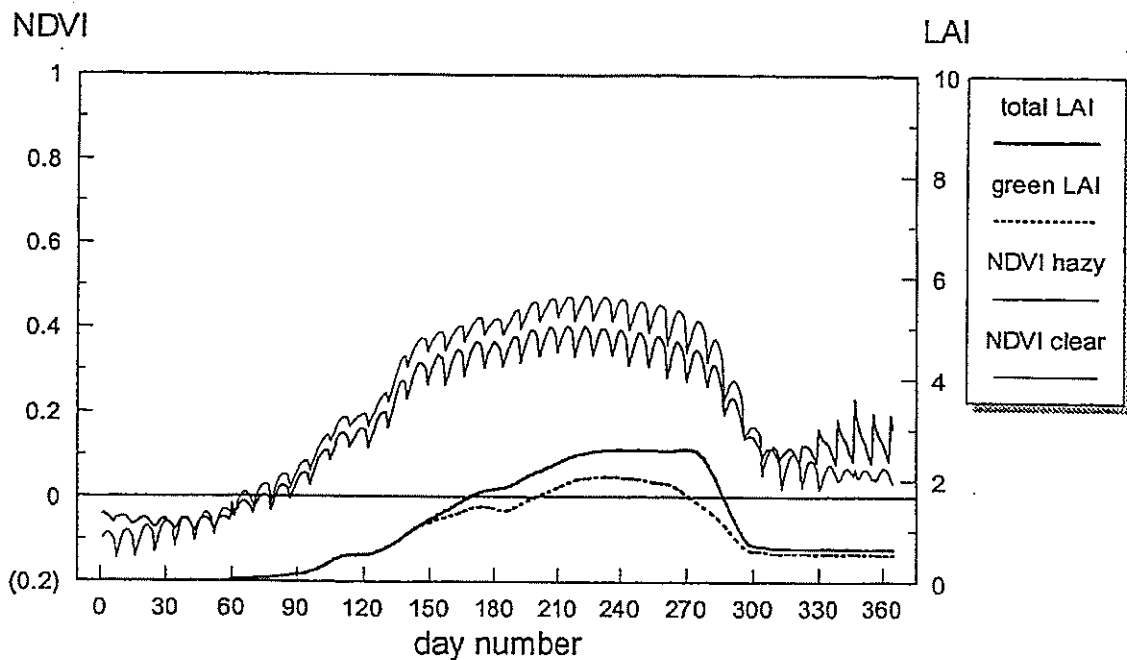


Fig. 11.1 NOAA simulated NDVI and LAI profiles for Flevoland area

green and total LAI for the crops potato, sugar beet and winter wheat. For the canopy reflectance modelling by means of the SAILH model, optical properties of green and yellow leaves of these crops in AVHRR bands 1 and 2, as well as the leaf angle distributions, were estimated from literature data. Actual NOAA orbital data were used to determine the daily viewing direction and solar zenith angle at the time of overpass. In the coupled SAILH - atmosphere model two extreme visibilities (5 km and 40 km, hazy and clear) were taken to compute the associate planetary reflectances in both bands for the three crops considered and for bare soil. By means of Landsat images the land use in the area had been estimated to consist of 21 % potato, 15 % sugar beet, 32 % winter wheat and 32 % bare soil. Using these proportions, weighted average planetary reflectances were determined for each day in both bands (red and near infrared) and finally the NDVI was calculated. Fig. 11.1 shows the simulated NDVI for clear and hazy atmospheric conditions and also the green and total LAI, which in this case are weighted averages over the land use classes.

It is demonstrated in Fig. 11.1 that the simulated NOAA-NDVI profile roughly resembles the profile of LAI, especially green LAI. The ripples in the NDVI signal are caused by the day-to-day variation of the viewing angle and (to a less extent) the solar time of day, which have a cycle period of



about 9 days for the NOAA satellite orbit. Near vertical viewing gives the highest NDVI in this case, which leads to the conclusion that the increase of NDVI on the ground for oblique viewing (more vegetation is observed) is more than compensated by the decrease due to atmospheric influence when viewing obliquely through the atmosphere.

From the months March to October the NDVI for a clear atmosphere is higher than for a hazy atmosphere, as expected, but from November to February the reverse is found in the simulated results. This phenomenon is known as the terminator-effect, i.e. the effect that artificially high NDVI values are found in the vicinity of the terminator, the line on the globe which marks the transition between the solar day and night. The explanation is that at low solar elevation the red light from the sun is attenuated more strongly in the atmosphere than the near infrared radiation, so that the target is illuminated by predominantly near infrared radiation, thus causing high NDVI values. The same effect, but then for visible light, is responsible for the apparent reddish colour of the sun just before sunset. In Fig. 11.1 the effect is stronger at the end of the year than at the beginning. This is caused by a gradual shift of the actual times of satellite overpass towards sunset time.

In the LAI profiles the rapid growth of winter wheat in April-May, potatoes in May-June and sugar beet in July are clearly recognized and this is also reflected in the NDVI profiles. The yellowing of these crops, expressed in the profile of green LAI, is accompanied by a declining NDVI, so one may conclude that the NDVI really is an indicative measure of green vegetation development.

A comparison of the simulated NDVI profiles of Fig. 11.1 with the actual NOAA-11 NDVI profile of the same area after temporal smoothing (Van Dijk *et al.*, 1994) is shown in Fig. 11.2. From this it is concluded that both agree in dynamic range and roughly in shape, but substantial differences of about 0.2 NDVI units also exist. The main reasons for these differences are probably inaccurate modelling of land use, crop emergence dates, leaf angle distribution, yellowing and harvesting activities on the one hand and too frequent (possibly sub-pixel) cloud cover over the Netherlands on the other. The absence of the terminator-effect in the actual NDVI profile in the winter period may be attributed to the fact that the atmosphere model is based on plane-parallel layers and a flat earth, which tends to exaggerate effects of a low solar elevation with respect to those for the actual spherical earth.

The above application of radiative transfer modelling is of the explanatory

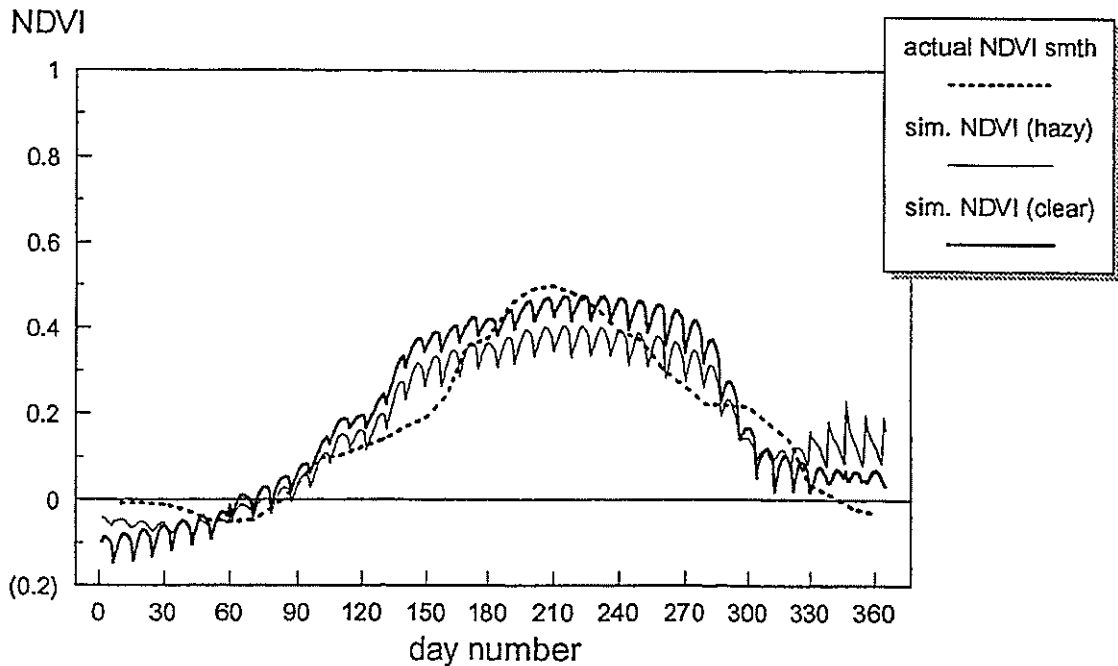


Fig. 11.2 Comparison of simulated and actual NDVI profiles

category, as it illustrates the behaviour of the NDVI under various influences and a comparison with actual NDVI data is made.

11.2 Red - near infrared feature space diagrams

In order to illustrate the effects of LAI, leaf angle distribution and soil brightness on the NDVI and other vegetation indices, the SAILH model has been applied to generate bidirectional reflectances in the red and the near infrared. The results have been plotted in a so-called feature space diagram, showing combinations of red and near infrared reflectance as points in a plane, with red on the X-axis and near infrared on the Y-axis. In such a diagram all points representing a certain NDVI lie on a straight line through the origin.

In Fig. 11.3 all combinations of ten LAI values plus zero LAI, two bare soils of different brightness and two leaf angle distributions are shown

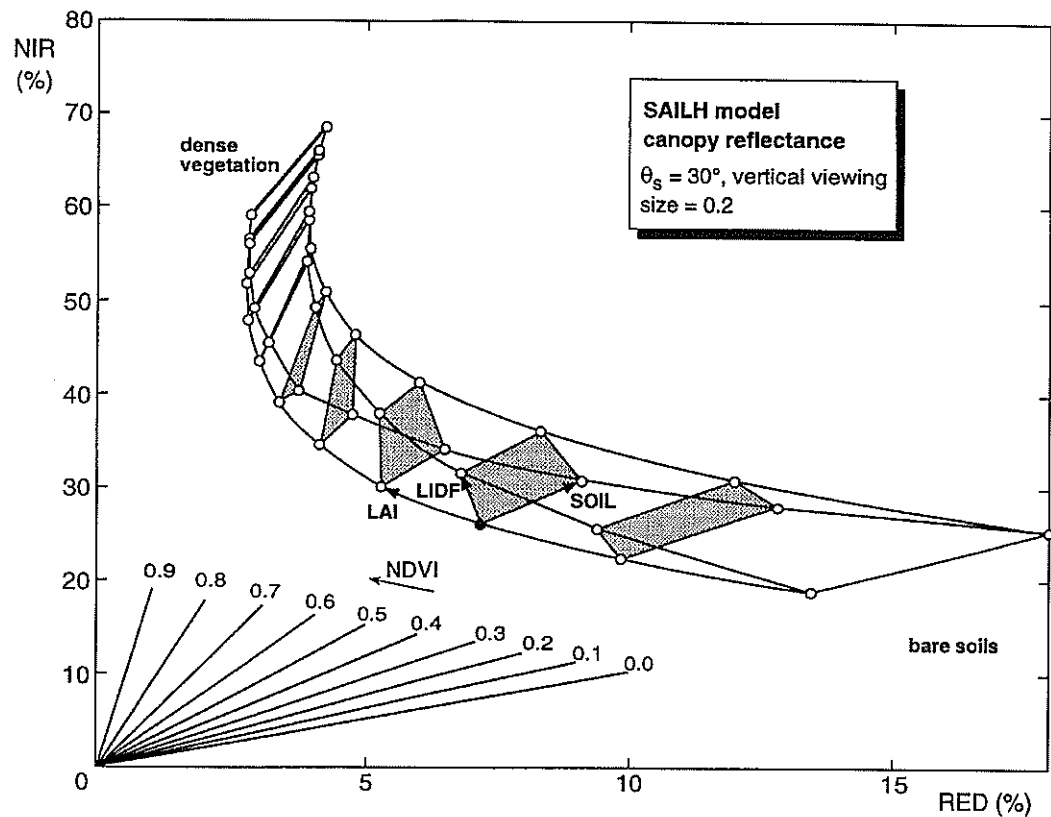


Fig. 11.3 Red - near infrared feature space diagram of canopy reflectance, with variations of LAI, LIDF and soil brightness

together with an NDVI scale indicated by a bundle of straight lines through the origin. The solar zenith angle is 45° and the hot spot size parameter is 0.2 in this case. The viewing direction is vertical. The LAI runs from zero to 8 in such a way that approximately equal steps in the near infrared reflectance are obtained, which is achieved by equidistant sampling of the function $\exp(-0.2 \text{ LAI})$. The quadrangular greyed shapes are areas of constant LAI. The leaf angle distributions are representative for sugar beet, namely plagiophile with 45° average leaf inclination (LIDF 4 of chapter 7) and plagio-erectophile with 65° average leaf inclination (LIDF 7). The soil brightness was varied by lowering the reflectances in both bands to 75 % of the bright soil values. The line through the bright and darker bare soil points also runs through the origin, which means that the NDVI is insensitive to soil brightness differences as long as there is no vegetation on it. From Fig. 11.3 it can be concluded that for soils covered with vegetation darker soils produce a slightly higher NDVI, but at very high LAI the soil influence disappears. The leaf angle distribution also has an effect on the NDVI. For



low to moderate LAI the effect is such that less steep leaves produce a higher NDVI, which is caused by the greater leaf area projection in vertical direction, so that more vegetation is observed. For high LAI a leaf angle distribution with less steep leaves produces higher reflectances in both bands, but the relative effect of this on the red is greater, so that the NDVI decreases. There is only one LAI where (in this case) neither the leaf slope nor the soil brightness have a significant influence on the NDVI and this is at $\text{LAI} \approx 3$. The relationships between LAI and NDVI are such that an increase of LAI always gives a higher NDVI, so that the NDVI is a good indicative measure of the amount of (green) leaf area, but the relationships are highly non-linear.

Based on similar simulations with the SAIL model, several other vegetation indices have been proposed, e.g. WDV (weighted difference vegetation index) by Clevers (1989), SAVI (soil adjusted vegetation index) by Huete (1988), and TSAVI (transformed soil adjusted vegetation index) by Baret *et al.* (1989). Of these, SAVI and TSAVI are still normalized differences of near infrared and red reflectance, but with a shifted origin in order to better correct for soil variations at moderate LAI. The WDV however measures the vertical distance from the soil line and in Fig. 11.3 this means that lines parallel to the soil line are lines of constant WDV. From the diagram it can be concluded that at moderate LAI the WDV increases slightly with soil brightness and that the leaf slope has a considerable effect on the WDV. However, Clevers (1986) found that for a constant leaf angle distribution the relationship between WDV and LAI can be approximated by a simple exponential function - this is known as the CLAIR model (Clevers Leaf Area Index by Reflectance) -, which makes the WDV suitable for direct estimation of the LAI. The LAI is an important biophysical parameter for the monitoring of crop growth, as it largely determines the rate of photosynthesis. The exponential relationship between LAI and WDV has generally been confirmed by large numbers of *in situ* measurements of the spectral directional reflectance and the LAI for the crops potato, sugar beet, barley, oats and wheat (Bouman *et al.*, 1992), performed in the period 1980-1990. They found that the radiometric estimation of LAI has an accuracy comparable to that of the traditional (destructive and time-consuming) method (10 to 15 percent relative error).

The above application of canopy reflectance modelling belongs to the category prediction. From simulations with the SAIL model it was predicted that the WDV (under certain conditions) would enable estimation of the



LAI, and this was largely confirmed by measurements in the field.

11.3 Sugar beet yield forecasting experiment

The BCRS-project "Remote sensing assisted sugar beet yield forecasting" (Verhoef *et al.*, 1996) has been carried out by NLR in cooperation with Wageningen Agricultural University, the Institute for Rational Sugar production IRS and the Dutch sugar industry in order to investigate the possibilities of using optical satellite images in support of the conventional method of yield forecasting, which is based on field samples. In this project optical data from the satellites SPOT and Landsat (Thematic Mapper) of about 60 sugar beet fields in the province of Flevoland and the associate agronomical data (including final yields), provided by the sugar industry, were statistically analyzed. One of the Landsat images was recorded on 1 July 1994 under good atmospheric conditions and also still a considerable spectral variation amongst the sugar beet fields was present. It was expected that under these circumstances a reasonably good correlation between spectral data and final yield figures could be obtained, especially since according to crop growth modelling results, with the given range of sowing dates in 1994, the LAI on this day was estimated to vary from 1.0 to 2.5, so that the relation between WDV and LAI could be approximated as being linear. However, no correlations of any significance were found, due to the following possible causes:

- The relation between LAI on certain dates in the growing season and final yield might be less firm than predicted by crop growth modelling, in particular for sugar beet.
- Final yield figures might be inaccurate.
- LAI estimation from satellite data via the WDV might be inaccurate.

Evidence for the first cause is given by the fact that virtually no correlation has been found between the sowing date and the final yield, neither in 1993, nor in 1994. In a crop growth model the sowing date, or rather the emergence date, is of crucial importance for the final yield. Also, about 50 percent of the final yield variation could be explained by a cultivar effect, which indicates that this alone is much more important than the sowing date. However, also within individual (frequently used) cultivars no correlation



between final yield and WDVl was found.

The second possible cause is unlikely, since in that case also the above mentioned cultivar effect would not have been found. However, there is an effect of the harvesting date on final yield, since on average harvesting in December gives higher yields than in September. This effect was estimated to be responsible for about 30 percent of the final yield variation.

The last possible cause has been investigated by means of simulations with the OSCAR model of chapter 10. For the series of LAI values 1.0, 1.5, 2.0 and 2.5 normalized vertically upward radiances in Thematic Mapper bands 3 and 4 (red and near infrared) have been computed for varying soil brightness, leaf angle distribution, leaf colour, observation height, atmospheric visibility and terrain surrounding the target. Some of the results are presented in Figures 11.4 to 11.7 by means of red - near infrared feature space diagrams. Fig. 11.4 shows the effects of soil brightness, leaf slope and leaf colour on the relation between LAI and the location in the diagram for observations on ground level and under a clear atmosphere (visibility 40 km). The overlapping four shaded polygons enclose the areas in the feature space corresponding to LAI and soil brightness variation only. The soil brightness varies by a factor of about two, which is representative for the region studied. The leaf slope variation was created by taking LIDFs 4 and 7 of chapter 7 (45° and 65° average leaf slope, respectively), which is reasonably representative for sugar beet. Leaf colour was varied by taking green and yellow leaves, which is probably too extreme for sugar beet. From Fig. 11.4 one may conclude that already at ground level the relation between LAI and WDVl can seriously be influenced by other factors. However, when leaf slope and leaf colour variations are small, the WDVl is a good measure of LAI, since soil brightness variation is automatically corrected to a large extent. Taking the upper left polygon of Fig. 11.4 (green leaves and leaf slope 45°) the observation height was varied from ground level to satellite level, with two intermediate heights of 1 km and 5 km. The result is displayed in Fig. 11.5 for a background consisting of bright bare soil. The main effect is a movement of the polygon to the right in the feature space diagram, accompanied by some shrinkage. The movement is caused by scattering in the atmosphere, particularly in TM3 (red). The shrinkage is due to the limited transmission of contrasts (target reflectance differences) through the atmosphere. Fig. 11.6 shows the influence of the atmospheric visibility for observations at satellite height. Again the polygon moves to the right and shrinks (due to increased scattering and reduced transmission) when

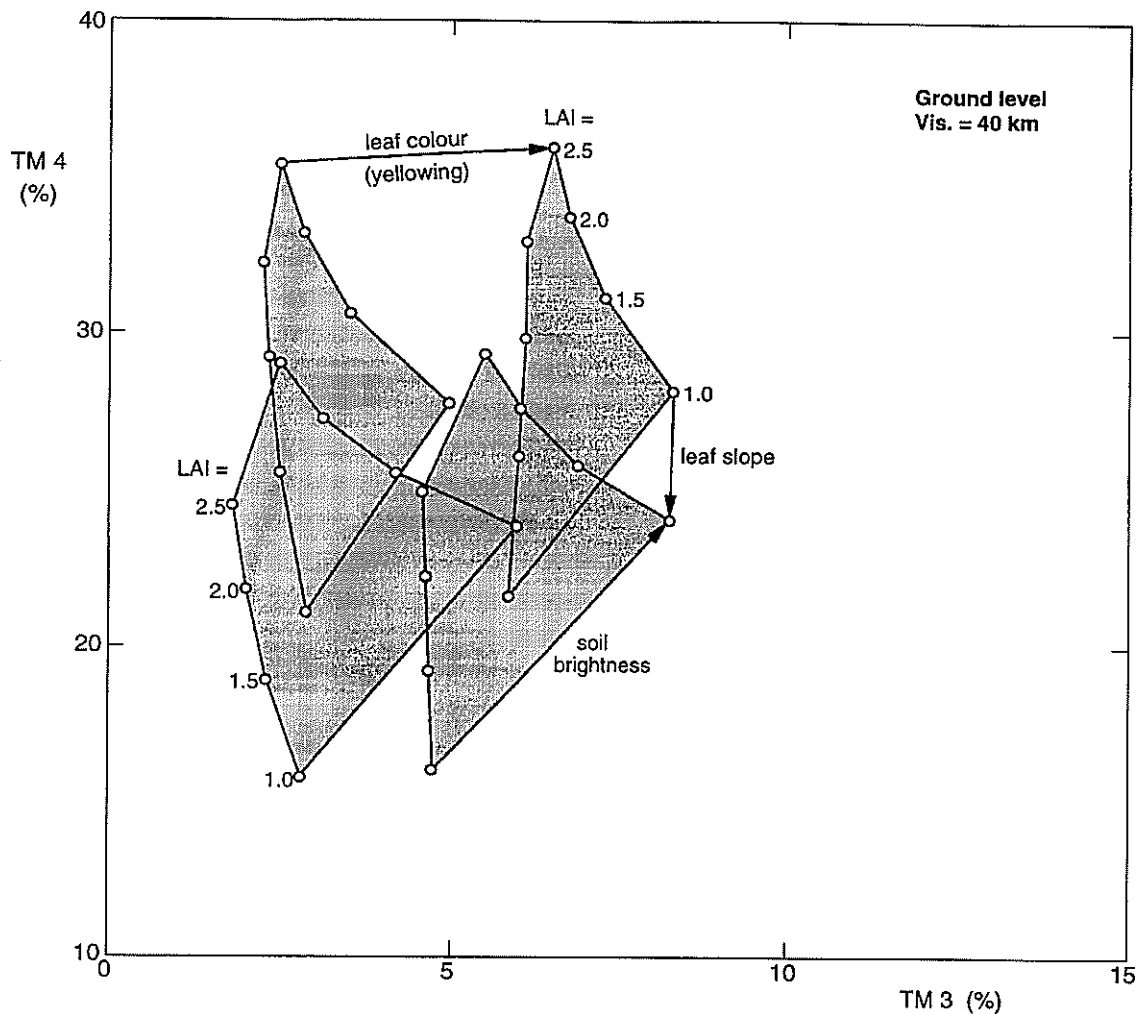


Fig. 11.4 Influences of soil brightness, LIDF and leaf colour on the relation between LAI and the red - near infrared normalized radiances on ground level

the visibility decreases. However, the most annoying effect is shown in Fig. 11.7. Here the background (surroundings of the target) is varied for the satellite altitude and a visibility of 10 km. Four combinations of background were taken: bright and dark soil, with and without vegetation on it. From this diagram it is concluded that the background has a considerable effect on the relation between LAI and WDV, if no location-dependent atmospheric correction is applied which takes this effect into account. These simulation results are consistent with regional trends, observed in the spectral satellite data from the sugar beet fields, that could be associated with areas having brighter bare soils or different land use in the neighbourhood. In the project

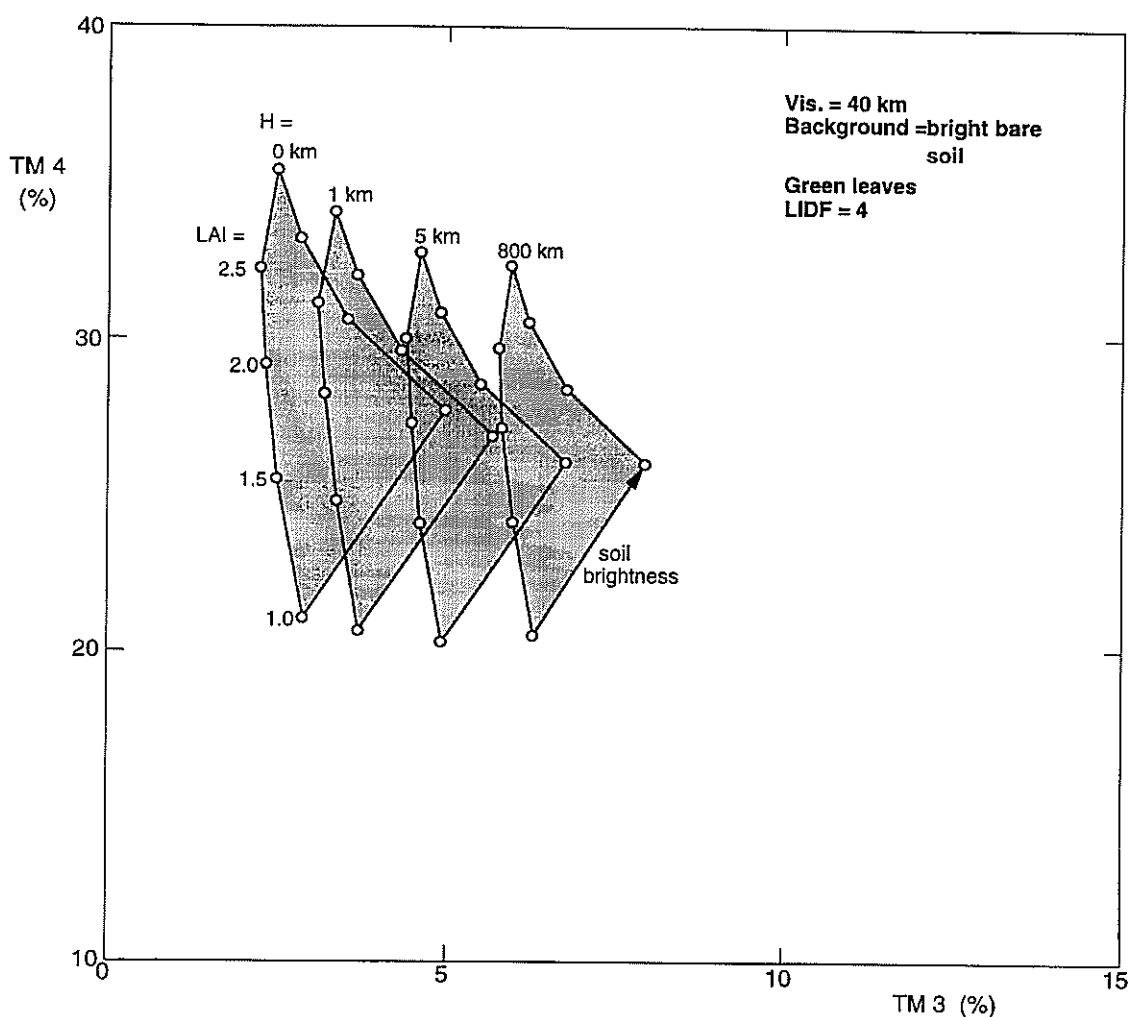


Fig. 11.5 Influence of observation height for 40 km visibility

(Verhoef *et al.*, 1996) such trends were removed by means of low-pass spatial filtering, but this is still far from ideal, since real existing regional spectral differences in the sugar beet crop might have been removed too in this way.

The above example is of the explanatory type and shows that radiative transfer modelling can give more insight in various remote sensing problems, and although no direct solutions of these problems are obtained, directions towards further progress can be deduced from the model simulation results.

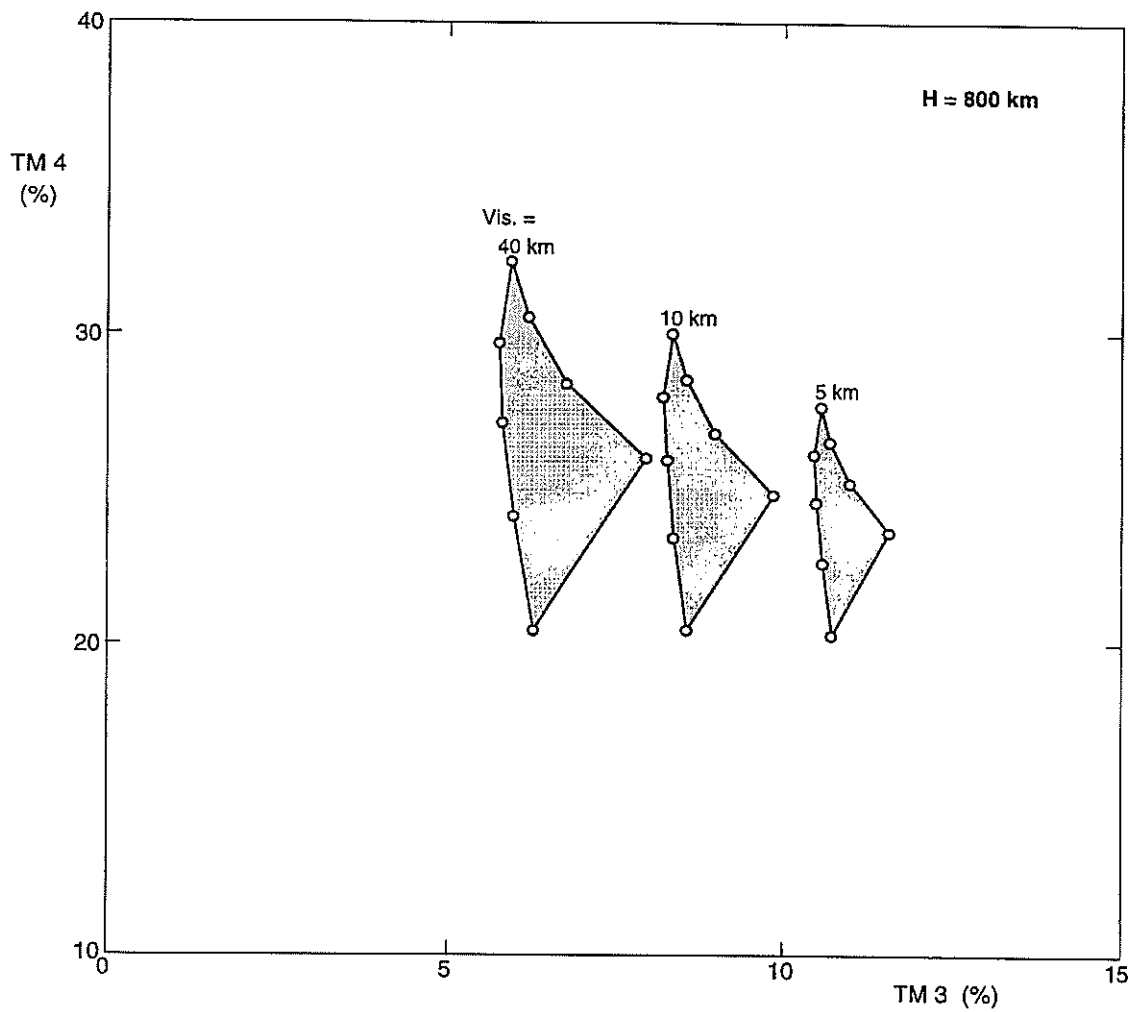


Fig. 11.6 Influence of visibility at satellite altitude (800 km)

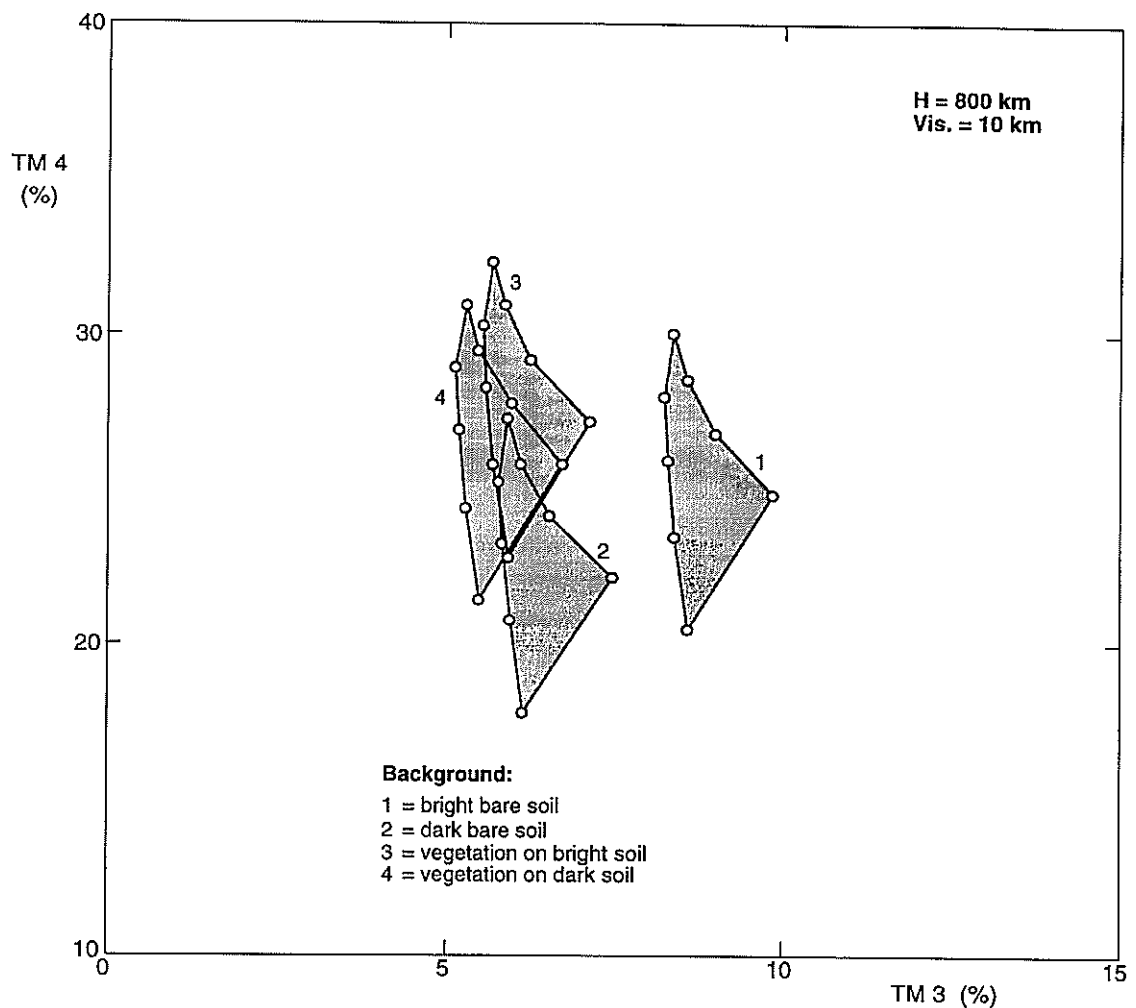


Fig. 11.7 Influence of background at satellite height and 10 km visibility

11.4 Model inversion

Numerical experiments on inversion of the SAIL model (Verhoef, 1984) have already been carried out (Goel & Thompson, 1984) before it was even published. These experiments were aimed at trying to deduce LAI, two leaf angle distribution parameters, leaf reflectance and transmittance, the fraction diffuse sky irradiance and the soil's reflectance from a large set of directional reflectance measurements in the near infrared, in other words from a monospectral but multi-angular data set. Several variations, mainly with respect to which parameters were allowed to vary and which were kept



constant, were also tried. An interesting conclusion was that the SAIL model, like the Suits model, is "totally invertible", meaning that when for a certain set of input parameters model-generated directional reflectance data are used as input for the inversion, the correct input parameters are retrieved. This means that each combination of input parameters leads to a unique angular pattern of directional reflectances. By adding different levels of noise to the directional reflectance data one is also able to investigate the influence of the radiometric resolution of the sensor on the results of the model inversion.

In practice, however, multi-angular reflectance data are not often available, because (quasi-)simultaneous measurement of the radiance in different directions from the same target is almost impossible. Therefore, it seems more appropriate to consider methods of model inversion that are based on multispectral radiance or reflectance data in one or only a few directions, which is more representative for situations found in practice.

A fundamental problem with model inversion based on multispectral data is that the number of unknowns is greater than the number of measurement data. For instance, for inversion of the SAILH model in a single spectral band the unknown parameters are:

- leaf reflectance
- leaf transmittance
- soil reflectance
- leaf area index
- leaf angle distribution
- hot spot size parameter

The measurement data in this case is given by the bidirectional reflectance for known directions of solar incidence and observation. If the leaf angle distribution is described by means of two parameters, then the number of unknowns is seven, of which three are spectral and four are related to the canopy structure. For multispectral measurements of the bidirectional reflectance in N spectral bands one obtains $3N+4$ unknowns versus N knowns, so there does not seem to be any point in applying model inversion based on multispectral data. However, if the optical properties of leaves and soil were known, then in principle it would be possible to deduce the four canopy structure parameters by means of model inversion from spectral reflectance measurement data in at least four spectral bands.



In many cases the optical properties of leaves and soil are not known, but this is not the same as complete ignorance, since there is much general knowledge on the spectral behaviour of leaves and soils, and this knowledge could be applied in the model inversion procedure. For instance, as shown by Jacquemoud & Baret (1990), entire reflectance and transmittance spectra of single leaves over the spectral range 0.4 - 2.5 μm can be predicted from only three parameters, namely chlorophyll concentration, water content and a parameter describing the leaf mesophyll structure. Similarly, soil reflectance spectra can be described by means of only one or two parameters. Incorporating this knowledge into the model inversion procedure would involve the estimation of eight or nine parameters, but with the advent of imaging spectrometers, which measure the radiance in 200 or more spectral bands, this could be feasible.

Another way of dealing with the problem of too many unknowns is that one simply makes guesses about the values of certain parameters, thereby accepting the possible errors of model inversion results caused by a mismatch of guessed and actual values. In fact, the use of vegetation indices like NDVI and WdVI is also based on certain assumptions, for instance that leaves are green and soils differ only in brightness and are spectrally nearly "grey". The use of the CLAIR model (Clevers, 1989) for estimation of LAI from the WdVI actually is also based on the assumption that all parameters except LAI and soil brightness are constant, and since the soil brightness effect is to a large extent corrected in the WdVI, the LAI remains as the only parameter to be estimated by means of inversion of this simplified model. An advantage of a simple semi-empirical model like CLAIR is that its inversion is very easy. However, a disadvantage is that it has to be recalibrated for each new crop type and (possibly) for each different measurement condition. Furthermore, when changes in LAI are accompanied by systematic changes in other parameters like leaf slope and leaf colour, the near-exponential relationship between WdVI and LAI does no longer hold, so that application of the CLAIR model becomes questionable in that case.

Operational application of model inversion is hampered by the fact that (classical) inversion of a model like SAILH requires many iterations. Even when a SAILH model calculation takes only less than a millisecond, several thousands of iterations might be required to reach a solution, so that application of this technique to each pixel of a multispectral image is out of the question. The use of Artificial Neural Networks (ANN) might be a good alternative in this respect. These are able to perform a non-linear mul-



tidimensional mapping of one space to another and this is exactly what is needed for model inversion from multispectral reflectance data. Neural networks are trained by learning from examples and when used for model inversion the most obvious training material consists of the results of many (forward) model calculations for a set of representative cases. In this way the knowledge of radiative transfer in leaf canopies is "recorded" in the ANN. Once sufficiently trained, the ANN would be able to process large amounts of multispectral data at high speed.

A trivial example of a neural network is a single neuron with two input weights and a linear output. Such a "network" would already be able to realize the WDVI function if reflectances in the red and near infrared were used as input. In general, each neuron with its associated input weights represents a hyperplane in the multispectral data space, and the output of the neuron is related to the perpendicular distance from this hyperplane, which thus acts as a decision boundary. Adding more neurons with different input weights allows to achieve a better localization in the feature space, and therefore also a better estimation of canopy parameters. In this way, neural networks can be imagined to bridge the gap between the simple WDVI approach and classical multidimensional model inversion methods.

11.5 Estimation of fraction APAR

Photosynthesis is the process by which plants convert solar energy into chemical energy (CO_2 assimilation into carbohydrates). Visible radiation from the sun and the sky is absorbed by leaf chlorophyll and a fraction of the absorbed radiant energy is used for photosynthesis, one of the main drivers of plant growth, besides the availability of water and nutrients. Chlorophyll is responsible for the green colour of plants, which is caused by strong absorption of blue and red, and a somewhat less strong absorption of green light. APAR is the absorbed photosynthetically active radiation of a vegetation canopy and thus is a measure of the growth potential of a crop. The fraction APAR or FPAR is the fraction of the incident visible radiation that is absorbed inside the canopy. This quantity can easily be calculated by means of the SAIL model, since the hemispherical reflectances and transmittances of the leaf canopy layer, which are required for this calculation, are already part of the output of the SAIL subroutine.

In Clevers *et al.* (1994) the PROSPECT leaf model (Jacquemoud & Baret,

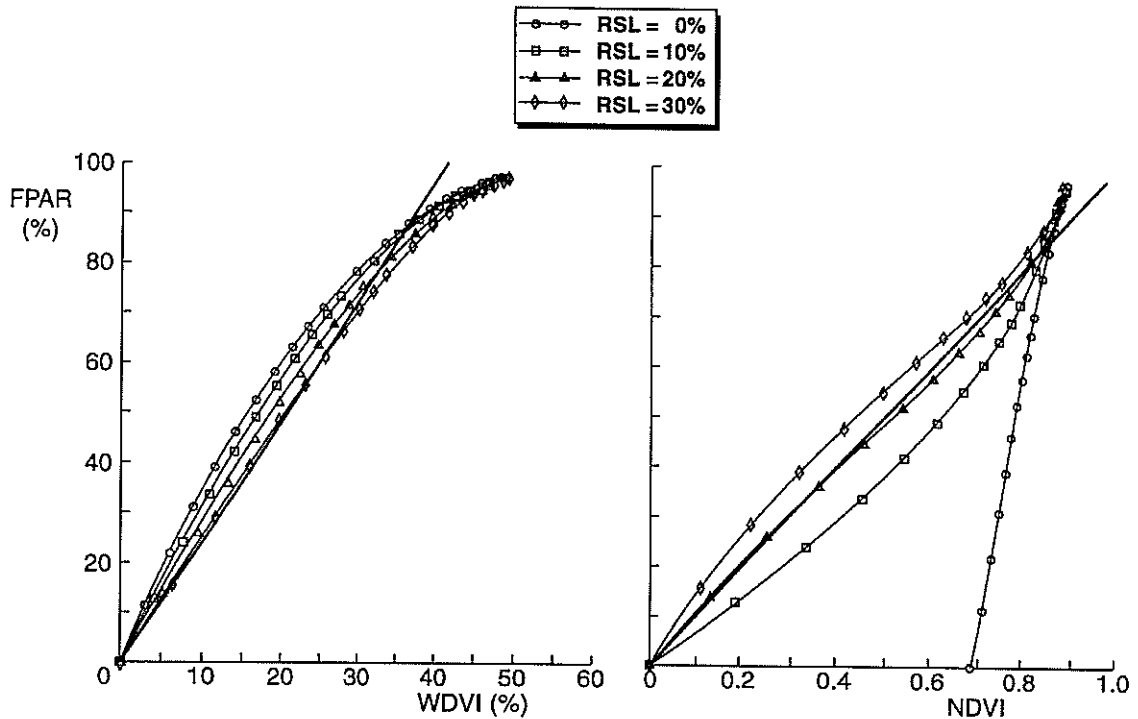


Fig. 11.8 Relationships between FPAR and vegetation indices as simulated by a combined PROSPECT-SAIL model (after Clevers et al., 1994)

1990) was combined with the canopy model SAIL in order to simulate the relationships between FPAR and the vegetation indices NDVI and WDVI under various circumstances. The conclusion of this study was that FPAR can be estimated fairly well (accuracy of the order of 10 percent) from a linear relationship with either the NDVI or the WDVI. See Fig. 11.8. Here RSL is the soil's reflectance (assumed equal in both bands) and the variation in NDVI, WDVI and FPAR was generated by means of LAI variation for otherwise "standard" input conditions. For $RSL = 0$ the NDVI is high for any LAI because this is a relative index of near infrared reflectance, which for a black soil is mainly determined by leaf optical properties.

11.6 Albedo estimation

In studies of the earth's radiation budget and of possible climate changes related to the enforced greenhouse effect the surface albedo plays an import-



ant role. The albedo a is defined as the ratio of reflected and incident shortwave (0.4 - 2.5 μm) radiation. Estimation of the albedo for vegetated surfaces from satellite observations is complicated by the fact that this requires integration of the directional reflectance over the exitent hemisphere and over the solar spectrum, whereas most optical remote sensing instruments measure only the radiance in one direction and in a limited number of spectral bands. Imaging spectrometers will be able to cover the shortwave range of the spectrum with high spectral resolution, but the number of different viewing directions remains very small. It is therefore almost inevitable that for the estimation of the surface albedo from remote sensing observations models of the bidirectional spectral reflectance be applied. By means of these models one could find relations between sparsely sampled spectral directional reflectance observations and the associate albedo. As satellite instruments measure the planetary (top-of-atmosphere) spectral radiance, not the surface reflectance, it will be necessary to include in these models also the atmospheric effects. The model OSCAR, described in chapter 10, would be a good candidate for the study of the above mentioned problems, as it includes the bidirectional reflectance of vegetation surfaces as well as scattering and absorption of radiation in the atmosphere. Furthermore, this model is fast (more than one thousand full model calculations per second), which is certainly an advantage, given the large numbers of simulations that would be required to perform the mentioned integrations over the upper hemisphere and the solar spectrum. Radiative transfer modelling can be applied here to explore different strategies for estimating the albedo, such as by direct extrapolation from observations under a limited number of viewing directions and spectral bands, or by using a model inversion step followed by forward modelling over the entire upper hemisphere. Also the dependencies of the surface albedo on the solar elevation angle and the atmospheric conditions can easily be investigated by means of an integrated model, whereas these would be very hard to obtain from field measurements or from remote sensing imagery, in particular because these factors cannot be controlled independently from one another.



12 CONCLUDING REMARKS

Radiative transfer models are important tools for the interpretation of optical remote sensing data. By means of these models one is able to explain the relationships between physical object properties and observational conditions on one hand and quantitative radiometric values extracted from remotely sensed images on the other.

The radiative transfer models discussed in this thesis consider the extinction, scattering and absorption of electromagnetic radiation in plane-parallel layers of infinite lateral extension (one-dimensional models). Radiative transfer in vegetation canopies and in the atmosphere can be modelled in this way, although one should always realize that this is an abstraction of reality, as in the real world both types of object have a definite 3D-structure.

Models of the bidirectional reflectance of vegetation canopies and of radiative transfer in the atmosphere both include a surface reflectance component at the bottom of the layer: in the former case this is the soil background and in the latter it is the earth's surface reflectance. In both cases the effective reflectance at the top of the layer can be derived by means of the adding method of Van de Hulst (1980). This method of coupling is simple and effective, and allows a clear identification of the contributions from the background surface and the layer above it. Application of this method requires the determination of all reflectance and transmittance quantities of the plane-parallel layer, but the advantage is that these quantities are exclusively related to the optical properties of the material of which it consists. The adding method can also be applied to an ensemble of many plane-parallel layers, bounded by a reflecting surface at the bottom. In this way it is possible to derive the planetary reflectance of the earth-atmosphere system by combining the soil with one or more vegetation layers and one or more atmospheric layers on top of it. Also the so-called adjacency effect can easily be expressed in the representation which is used in the adding method (cf. chapter 10). However, in order to enable application of the adding method to the vegetation-atmosphere system, the interaction of the radiant fluxes at the interface between both types of medium must be described by means of reflectance quantities which are compatible with both. This is considerably simplified if both media can be described by means of a similar radiative transfer model.

Radiative transfer in the atmosphere is characterized by a highly anisotropic scattering phase function, due to the strong forward scattering of aerosols.



On the other hand, because of the random orientation of the particles, the interception of incident flux is independent of the incidence direction. Also, scattering depends only on the scattering angle (the angle between the directions of incidence and exitence). These are properties of a so-called isotropic medium, and they allow to describe radiative transfer more effectively by means of spherical harmonic functions, which has led to special solution methods adapted to this case.

Radiative transfer in leaf canopies is characterized by a less anisotropic scattering than is present in the atmosphere, but otherwise it is more complex, because the interception of incident flux does depend on the incidence zenith angle and scattering is essentially bidirectional in nature. With respect to the scattering behaviour, the simplest one can think of for vegetation is the one obtained for a so-called azimuthally isotropic leaf canopy, which results if one assumes that the leaves have random leaf azimuth and are randomly positioned in the layer. There is only one type of leaf canopy that can be treated in the same way as the atmosphere, namely the canopy with a spherical leaf angle distribution, but this is too special to be of general interest. Therefore it can be concluded that in general vegetation canopies cannot be modelled by means of the theory specifically developed for atmospheres.

Four-stream radiative transfer theory, which is based on the differential equations of the Suits and SAIL vegetation bidirectional reflectance models, can also be applied to the atmosphere, as the atmosphere is an isotropic medium, which means that it is just a special case of the broader class of azimuthally isotropic media. However, the strong anisotropy of aerosol scattering, in combination with the weak absorption, leads to a considerable numerical inaccuracy of a four-stream atmosphere model. The reason for this is that in a four-stream model the diffuse radiation is represented by means of only two hemispherically isotropic fluxes. This representation is too coarse for an accurate description of aerosol scattering.

The generalization of four-stream theory to an $(N+2)$ -stream model, which can successfully be applied to vegetation canopies as well as to atmospheres, has been formulated by defining N diffuse fluxes on the basis of an equal-weight tessellation of the sphere of directions. Similarly to the matrix-vector generalization of two-stream Kubelka-Munk theory by Chen (1985), this too leads to a matrix-vector formulation of the differential equations. Similar systems of differential equations also occur in the theory of the discrete ordinates method (DOM).



What distinguishes the $(N+2)$ -stream theory from other theories is the equal-weight tessellation and the incorporation of both a specular flux from the sun and a radiance in the direction of the observer in the system of differential equations. The equal-weight tessellation leads to a uniform reflectance matrix in case the reflection of a surface is Lambertian. The incorporation of both directions (incidence and exitence) allows formulating the analytical solution by means of expressions in which symmetry with respect to these directions can be recognized immediately. This leads to formulas which inherently obey the reciprocity relations.

In the analytical solution of the $(N+2)$ -stream system of differential equations the linear transformations applied in order to reduce the computational complexity are similar to those usually applied in DOM-theory: decoupling of zenithal and azimuthal dependencies by means of a cosine transform in the azimuthal domain and eigenvector decomposition in the remaining zenith domain. The general solution consists of linear combinations of exponential functions. Substitution of the boundary conditions leads to expressions for the reflectance and transmittance of a layer. Special attention has been paid to possible numerical problems associated with the occurrence of equal eigenvalues. The analytical solution of the $(N+2)$ -stream model now has been formulated in such a way that direct implementation on a computer is possible without introducing numerical instability. The result is a model that can be applied for a wide range of input conditions.

Numerical tests with the $(N+2)$ -stream model for aerosol and Rayleigh scattering show that the model can be applied for all optical thicknesses occurring in practice. Tests for extremely high optical thickness and zero absorption (conservative scattering) indicate no significant numerical deficiencies and in all cases the reciprocity relations were verified.

The modelled dome-shaped directional transmittance pattern occurring at very high optical thickness very well matches an existing empirical relation for the radiance distribution of a uniform overcast sky. This pattern may be applied also to the directional transmittance of single leaves, as the optical thickness of single leaves is also very high.

By means of the $(N+2)$ -stream model it could be established that a four-stream approximation for an atmosphere containing common amounts of aerosol particles gives considerable inaccuracies in the directional patterns of atmospheric reflectance and transmittance. In general the results for the four-stream model are less anisotropic than for the $(N+2)$ -stream model. This is explained by the fact that the hemispherical isotropic fluxes of the four-



stream model lead to a faster damping of directional variations in the multiple scattering term than is the case in the $(N+2)$ -stream model. It is recommended that a similar comparison of performances be carried out for vegetation canopies in the near infrared, where multiple scattering also forms an important contribution.

The SAIL vegetation canopy reflectance model has been redescribed with emphasis on the extinction and scattering coefficients and their dependence on the leaf angle distribution. A parameterization of the leaf angle distribution by means of two quantities related to the average leaf inclination and the bimodality of the distribution can mimic most common types (planophile, erectophile, plagiophile and extremophile) very well. The coefficients of extinction and bidirectional scattering of the SAIL model might directly be applied in an $(N+2)$ -stream version of this model, leading to a higher accuracy in the multiple scattering contribution, which plays an important role in the near infrared.

However, implementation of the hot spot effect is considered more important, as this effect has a substantial influence in all parts of the spectrum, and in a wide region around the actual hot spot. The hot spot effect has been incorporated into the SAIL model by means of the theory of Kuusk. The improved model is called SAILH. In the manner it was implemented, it only modifies the single scattering contribution to the bidirectional reflectance of a leaf canopy. This means that it might be incorporated in an $(N+2)$ -stream version of SAIL with comparable ease. However, even such a model still has its limitations, since e.g. crop rows, clumpiness, leaf gloss and polarization are known to have considerable effect on the bidirectional reflectance under certain conditions.

As a result of the hot spot effect, reflectance profiles as a function of time or LAI need reconsideration, as the growth strategy of the species has a strong influence on this profile. Especially crops with a constant leaf size to crop height ratio show a much more favourable LAI - reflectance relation than formerly obtained by not considering the hot spot effect. The near infrared reflectance for this type of growth strategy remains sensitive to the LAI when it has largely exceeded the biologically plausible maximum, thus indicating better prospects for LAI retrieval. The former limitation due to saturation of the reflectance at high LAI remains in existence only for crops which have a particular growth strategy, namely one for which the product of the LAI and the leaf-size-to-canopy-height ratio is constant.

A theoretical limitation of Kuusk's analysis of the hot spot effect has been



investigated by means of a more accurate calculation of the bidirectional projection overlap function for circular leaves. The results show that the error due to Kuusk's approximation in most common cases is small, while using the more exact calculation involves a hundredfold reduction of computation speed.

A four-stream atmosphere model has a limited numerical accuracy. However, it can still be used for the correction of optical remote sensing imagery, as the uncertainty about the input parameters usually will generate errors which are much greater than the numerical error of the model. When a four-stream atmosphere model is inverted for estimation of the aerosol optical thickness, all other parameters must be assumed constant. In that case the retrieved optical thickness values are probably inaccurate. However, our main concern is not the accuracy of the input parameters of the model, but rather the accurate estimation of the parameters describing the atmospheric effect on the images. As both the darkest pixel method and the cloud shadow method are applied by using multispectral image data and spectral relationships, the estimation of the correction parameters is probably still accurate enough for application in situations where no atmospheric *in situ* data are available, in spite of the inaccuracy of model input parameters. A better modelling of the adjacency effect is recommended however, since this effect can be unexpectedly strong in certain situations.

The OSCAR model, in which the SAILH model has been combined with a four-stream atmosphere model, also has a limited numerical accuracy, but again it can be stated that in many cases a high accuracy will not be required. For instance, when one is interested in studying certain systematic effects, the computed radiances might be somewhat in error, but the predicted trends probably are still realistic. In spite of its limited accuracy, the OSCAR model is able to simulate many physical aspects of optical remote sensing, including the hot spot effect, path radiance, atmospheric transmittance and the adjacency effect. A model like OSCAR can be very useful for educational purposes and can assist in the preparation for future earth observation missions featured by advanced optical instruments, like MERIS, AATSR and PRISM.

Applications of radiative transfer modelling in optical remote sensing fall in the categories prediction, explication and model inversion. Model predictions can be used to find relationships between object parameters and remote sensing observables or to find transformed variables which are maximally sensitive to a certain parameter and least sensitive to other parameters.



Explication by means of model simulations can be used to illustrate certain phenomena in actual remote sensing images. On the other hand, when some of the effects observed in the actual data cannot be explained by means of the model, this might be a reason for modifying the model. This feedback is important, although it should not be exaggerated, as it is equally important to keep a model as simple as is reasonably possible. In other words, the evidence that a model disagrees with reality must be very convincing before one should decide to modify the model.

In the context of radiative transfer models, model inversion is used to estimate input parameters of the model by means of observed remotely sensed data. When an input parameter of the model is an important object property, this can be a very useful means to extract quantitative information from remote sensing data. However, the field of model inversion is plagued by great practical difficulties, mostly due to the high number of unknown parameters. In addition, practical application of model inversion to remote sensing imagery is hampered by the substantial computational effort involved. A single model inversion requires several iterations on the forward model, and doing this for every pixel in the image seems out of the question. In this respect, artificial neural networks, trained by means of the data from a comprehensive set of forward model calculations, might be a more practical solution, also because probabilistic approaches dealing with the unknown parameters might be realized by training the network with representative input data sets in which the probability of occurrence of certain parameters is taken into account.

In the preparation for future earth observation missions, which will be featured by hyperspectral and multidirectional remote sensing data, much attention should be paid to how this wealth of information can effectively be used. Radiative transfer modelling may assist in this process, and since also co-registered data from the thermal infrared will be acquired by future sensors, extension of radiative transfer models towards this spectral region should be considered.

As to further developments in the field of radiative transfer modelling for applications in optical remote sensing over land surfaces it can be stated that in general there is still a large gap between theory and the practical application of these models. This is partly caused by the fact that most current models apply to laterally homogeneous atmospheres over uniform landscapes. Even a simple four-stream model like OSCAR shows (cf. chapter 10) that for instance the adjacency effect has such a strong influence on



spectral radiances detected by remote sensing instruments, that spectral signatures from surface objects are always contaminated by radiation from objects in the surroundings. Realistic modelling of this effect actually is only possible if one takes full account of the spatial heterogeneity of the landscape, which implies that a much closer connection between radiative transfer models and remotely sensed images should be established. Also the atmosphere itself often contains locally varying amounts of haze and water vapour, and probably a 3D-modelling of the atmospheric layer, "draped" over the image representing the landscape, might be the best direction for making progress in the understanding and interpretation of optical remote sensing imagery. This form of 3D scene radiation modelling could be combined with model inversion techniques in order to extract biophysical land surface parameters from the image data.

Another reason for this gap between theory and practice may be the fact that the relations between object parameters and remote sensing data are multidimensional and non-linear, which makes interpretation by humans complicated. Perhaps this is the main reason why e.g. vegetation indices, which reduce all spectral information to just one parameter, are still so popular. In this respect there is a task for modellers to apply more advanced presentation techniques in order to make the remote sensing user community aware of the information that is thrown away if one reduces everything to just one index.

From a scientific point of view, in the field of radiative transfer modelling there are still a few possible subjects needing further investigation:

- the effect of N on the numerical accuracy of the $(N+2)$ -stream model
- implementation of an $(N+2)$ -stream vegetation canopy reflectance model, including the hot spot effect
- the influence of non-Lambertian reflectance and transmittance of plant leaves and the soil, as well as the effects of heterogeneity due to clumpiness and crop rows
- polarization of atmospheric scattering and canopy reflectance
- integrated $(N+2)$ -stream modelling of atmospheric scattering and the earth's surface reflectance for heterogeneous landscapes

Besides that these subjects are of scientific interest, they may contribute as well to a more effective use of advanced optical remote sensing instruments.

REFERENCES

- Allen, W.A., Gayle and Richardson, 1970, "Plant canopy irradiance specified by the Duntley equations", *J. Opt. Soc. Am.* 60(3): 372-376.
- Baret, F., Guyot, G. and Major, D., 1989, "TSAVI: a vegetation index which minimizes soil brightness effects on LAI and APAR estimation", Proc. 12th Canadian Symp. on Remote Sensing, Vancouver, Canada, July 10-14, 1989.
- Bouman, B.A.M., Van Kasteren, H.W.J. and Uenk, D., 1992, "Standard relations to estimate ground cover and LAI of agricultural crops from reflectance measurements", *Eur. J. Agron.* 1(4): 249-262.
- Bunnik, N.J.J., 1978, "The multispectral reflectance of shortwave radiation in relation with their morphological and optical properties", Thesis, Agricultural University, Wageningen.
- Camillo, P., 1987, "A canopy reflectance model based on an analytical solution to the multiple scattering equation", *Rem. Sens. of Env.* 23: 453-478.
- Chen, J., 1985, "Kubelka-Munk equations in vector-matrix forms and the solution for bidirectional vegetative canopy reflectance", *Appl. Optics* 24(3): 376-382.
- Clevers, J.G.P.W., 1986, "Application of remote sensing to agricultural field trials", Thesis, Agricultural University, Wageningen.
- Clevers, J.G.P.W., 1989, "The application of a weighted infrared-red vegetation index for estimating Leaf Area Index by correcting for soil moisture", *Rem. Sens. of Env.* 29: 25-37.
- Clevers, J.G.P.W., Van Leeuwen, H.J.C. and Verhoef, W., 1994, "Estimating the fraction APAR by means of vegetation indices: A sensitivity analysis with a combined PROSPECT-SAIL model", *Rem. Sens. Reviews* 9: 203-220.

Colwell, R.N. (Editor-in-Chief), 1983, "Manual of Remote Sensing, 2nd edition", Am. Soc. of Photogrammetry, Falls Church, VA.

Cooper, K., Smith, J.A. and Pitts, D., 1982, "Reflectance of a vegetation canopy using the Adding method", *Appl. Optics* 21(21).

Deirmendjian, D., 1969, "Electromagnetic scattering on spherical polydispersions", American Elsevier Publ. Co., New York.

Van Dijk, A., Roozkrans, J.N., Berkhout, J.A.A., Verhoef, W., Nieuwenhuis, G.J.A., Mùcher, C.A., Meuldijk, D. and Verschoor, M., 1994, "Prefase: Spectral Analysis of Mixed Agricultural Sites (PRESAMAS) by means of crop growth and radiometric simulation to interpret NOAA-AVHRR imagery", BCRS Report 94-11, Netherlands Remote Sensing Board, Delft.

Den Dulk, J.A., 1989, "The Interpretation of Remote Sensing, a feasibility study", Thesis, Agricultural University, Wageningen.

Duntley, S.Q., 1942, "The optical properties of diffusing materials", *J. Opt. Soc. Am.* 32(2): 61-70.

Elterman, L., 1970, "Vertical attenuation model with eight surface meteorological ranges 2 to 13 km", AFCRL, Bedford, Mass., NTIS-AD 707 488.

Epema, G.F., 1990, "Determination of Planetary Reflectance for Landsat-5 Thematic Mapper Tapes Processed by Earthnet (Italy)", *ESA Journal*, Vol. 14.

Epema, G.F., 1992, "Spectral reflectance in the Tunisian desert", Thesis, Agricultural University, Wageningen.

Gerstl, S.A.W. and Zardecki, A., 1985, "Discrete-ordinates finite-element method for atmospheric radiative transfer and remote sensing", *Appl. Optics* 24(1): 81-93.

Gerstl, S.A.W., 1988, "The angular reflectance signature of the canopy hot spot in the optical regime", Proc. 4th Int. Colloquium on Spectral Signatures of Objects in Remote Sensing", Aussois, France, Jan. 1988, ESA SP-287, pp. 129-132.

Grace, J., 1971, "The directional distribution of light in natural and controlled environment conditions", *J. Appl. Ecol.* 8: 155-165.

Goel, N.S. and Strebel, D.E., 1984, "Simple beta distribution representation of leaf orientation in vegetation canopies", *Agron. J.* 76: 800-802.

Goel, N.S. and Thompson, R.L., 1984, "Inversion of vegetation canopy reflectance models. IV. Total inversion of the SAIL model", *Rem. Sens. of Env.* 15: 237-253.

Goel, N.S. and Grier, T., 1986, "Estimation of canopy parameters for inhomogeneous vegetation canopies from reflectance data I. Two-dimensional row canopy", *Int. J. Rem. Sens.* 7: 665-681.

Goel, N.S., 1988, "Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data", *Rem. Sens. Reviews* 4: 1-212.

Goudriaan, J., 1977, "Crop micrometeorology: a simulation study", Thesis, PUDOC, Wageningen.

De Haan, J.F., 1987, "Effects of aerosols on the brightness and polarization of cloudless planetary atmospheres", Thesis, Free University, Amsterdam.

Huete, A.R., 1988, "A soil adjusted vegetation index (SAVI)", *Rem. Sens. of Env.* 23: 213-232.

Van de Hulst, H.C., 1980, "Multiple Light Scattering", Academic, London.

Huygen, J., 1988, "Duet User Manual, Version 2.0", MARS project, DLO-Staring Centre, Wageningen.

- Jacquemoud, S. and Baret, F., 1990, "PROSPECT: a model of leaf optical properties spectra", *Rem. Sens. of Env.* 34: 75-91.
- Kasten, F. and Raschke, E., 1974, "Reflection and transmission terminology by analogy with scattering", *Appl. Optics* 4(13): 460.
- Kimes, D.S. and Kirchner, J.A., 1982, "Radiative transfer model for heterogeneous 3-D scenes", *Appl. Optics* 21: 4119-4129.
- Kubelka, P. and Munk, F., 1931, "Ein Beitrag zur Optik der Farbanstriche", *Ann. Techn. Phys.* 11: 593-610.
- Kuusk, A., 1985, "The hot spot effect of a uniform vegetative cover", *Sov. J. Rem. Sens.* 3: 645-658.
- Markham, B.L. and Barker, J.L., 1987, "Thematic Mapper bandpass solar exo-atmospheric irradiances", *Int. J. Rem. Sens.*, Vol. 8, No. 3, 517-523.
- Minnaert, M., 1940, "De natuurkunde van 't vrije veld", I.W.J. Thieme & Cie, Zutphen, pp. 226-230.
- Moon, P. and Spencer, D.E., 1942, "Illumination from a non-uniform sky", *Trans. Illum. Engin. Soc.*, New York, Vol. 37: 707-712.
- Myneni, R.B., Asrar, G. and Kanemasu, E.T., 1987, "Light scattering in plant canopies: the method of successive orders of scattering approximations (SOSA)", *Agr. Forest Meteor.* 39: 1-12.
- Nakajima, T. and King, M.D., 1992, "Asymptotic theory for optically thick layers: application to the discrete ordinates method", *Appl. Optics* 31(36): 7669-7683.
- Nicodemus, F.E., 1970, "Reflectance nomenclature and directional reflectance and emissivity", *Appl. Optics* 9(6): 1474-1475.
- Nilson, T. and Kuusk, A., 1989, "A reflectance model for the homogeneous plant canopy and its inversion", *Rem. Sens. of Env.* 27: 157-167.

Norman, J.M. and Welles, J.M., 1983, "Radiative transfer in an array of canopies", *Agron. J.* 75: 481-488.

Rosema, A., Verhoef, W., Schroote, J. and Snel, J.F.H., 1991, "Simulating fluorescence light-canopy interaction in support of laser-induced fluorescence measurements", *Rem. Sens. of Env.* 37: 117-130.

Ross, J. and Nilson, T., 1975, "Radiative exchange in plant canopies", in: *Heat and Mass Transfer in the Biosphere*.

Ross, J.K. and Marshak, A.L., 1984, "Calculation of the solar radiation reflection from plant cover using the Monte Carlo method", *J. Earth Research from Space* 5: 58-67.

Saunders, R.W., 1988, "The determination of broad band surface albedo from AVHRR visible and near infrared radiances", MET O 19 Branch Memorandum No. 96, Clarendon Lab., Oxford.

Stammes, P., 1989, "Light scattering properties of aerosols and the radiation inside a planetary atmosphere", Thesis, Free University, Amsterdam.

Stammes, K., Tsay, S.-C., Wiscombe, W. and Jayaweera, K., 1988, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media", *Appl. Optics* 27(12): 2505-2509.

Sturm, B., 1981, "The atmospheric correction of remotely sensed data and the quantitative determination of suspended matter in marine water surface layers", in: *Remote Sensing in Meteorology, Oceanography and Hydrology*, Ellis Horwood Ltd., Chichester, England.

Suits, G.H., 1972, "The calculation of the directional reflectance of a vegetative canopy", *Rem. Sens. of Env.* 2: 117-125.

Suits, G.H., 1983, "Extension of a uniform canopy reflectance model to include row effects", *Rem. Sens. of Env.* 13: 113-129.

Turner, R.E., 1973, "Radiative transfer in real atmospheres", Technical Report, Environmental Research Institute of Michigan, NASA-CR-140199.

Verhoef, W. and Bunnik, N.J.J., 1975, "A model study on the relations between crop characteristics and canopy reflectance", NIWARS Publ. No. 31, Delft.

Verhoef, W. and Bunnik, N.J.J., 1976, "The spectral directional reflectance of crops. 1. Consequences of non-Lambertian behaviour for automatic classification. 2. Measurements on wheat and simulations by means of a reflectance model for row crops", NIWARS Publ. No. 35, Delft.

Verhoef, W. and Bunnik, N.J.J., 1981, "Influence of crop geometry on multispectral reflectance determined by the use of canopy reflectance models", Proc. Int. Colloquium on Spectral Signatures of Objects in Remote Sensing, Avignon, France, Sept. 1981.

Verhoef, W., 1984, "Light scattering by leaf layers with application to canopy reflectance modelling: the SAIL model", *Rem. Sens. of Env.* 16: 125-141.

Verhoef, W., 1985a, "Earth observation modelling based on layer scattering matrices", *Rem. Sens. of Env.* 17: 165-178.

Verhoef, W., 1985b, "A scene radiation model based on four-stream radiative transfer theory", Proc. 3rd Int. Colloquium on Spectral Signatures of Objects in Remote Sensing, Les Arcs, France, Dec. 1985, ESA SP-247, pp. 143-150.

Verhoef, W., 1988, "Numerical accuracy of a two-stream radiative transfer model from comparison with an N-stream model", Proc. 4th Int. Colloquium on Spectral Signatures of Objects in Remote Sensing, Aussois, France, Jan. 1988, ESA SP-287, pp. 157-162.

Verhoef, W., 1990, "Informatie-extractie uit multispectrale beelden met behulp van stralingsinteractiemodellen", NLR TP 90243 L.



Verhoef, W., Van Leeuwen, H.J.C. and Clevers, J.G.P.W., 1996, "Remote sensing assisted sugar beet yield forecasting", BCRS Report 95-24, Netherlands Remote Sensing Board, Delft.

Veugen, L.M.M. and Van Stokkom, H.T.C., 1985, "An atmospheric correction method using Guzzi-spectroradiometer data", Proc. 3rd Int. Colloquium on Spectral Signatures of Objects in Remote Sensing, ESA SP-247, pp. 383-388.

Weinman, J.A. and Guetter, P.J., 1972, "Penetration of solar irradiance through the atmosphere and plant canopies", *J. Appl. Meteor.* 11: 136-140.

De Wit, C.T., 1965, "Photosynthesis of leaf canopies", Agr. Res. Report 663, PUDOC, Wageningen.



SUMMARY

In the past few decades, optical remote sensing techniques have been used in order to acquire information on the properties of the earth's surface materials. In most of these remote sensing techniques the sun is used as the primary source of radiation and the radiation reflected by the surface (land or ocean) is detected by an electronic sensor placed on board an aircraft or a satellite. As the reflected radiation depends on the optical properties of the surface material, it is possible to acquire information about the spatial distribution of these surface properties by means of scanning. In this way digital images are produced which can be studied and interpreted by means of a computer.

For a quantitative interpretation of remotely sensed images it is necessary to have knowledge of the relations between surface properties and the resulting reflectance and its dependence on the wavelength and the directions of view and of the sun at the time of image acquisition. Also, one has to take account of the influence of the atmosphere, as the radiation reflected by the surface is modified on its way to the sensor. Radiative transfer models of the atmosphere and of surface objects like vegetation canopies have been developed in order to gain more insight in these relations.

In this thesis the theory of the transfer of electromagnetic radiation in media such as the atmosphere and vegetation canopies is discussed. On the basis of this theory one can construct radiative transfer models which describe how radiation interacts with the material inside the medium. In this way one can compute what part of the incident radiation is reflected at the top of a layer, what part is transmitted and how much of the incident radiation is absorbed.

For the application to remote sensing problems the spectral directional surface reflectance is the quantity of interest, as this is the quantity (apart from the atmospheric influence) that is detected by optical remote sensing instruments. Therefore, much attention is paid in this thesis to this and various other types of reflectance quantities that can be defined.

In chapter 1 a brief introduction is given on optical remote sensing techniques. Optical remote sensing refers to the part of the electromagnetic spectrum where the sun is the primary source of radiation. In wavelengths, this is the region from about 400 nm to 2500 nm, thus including the visible, the near infrared and the shortwave infrared. In the optical region radiative transfer in remote sensing applications can be described by means of only two processes: absorption and scattering. The rationale of the use of



radiative transfer models for the quantitative interpretation of optical imagery is also briefly discussed here.

Chapter 2 reviews the main radiometric quantities involved in radiative transfer models, such as radiant flux, intensity, irradiance, radiance, reflectance, transmittance and the coefficients of extinction and scattering. Also the radiative transfer equation is introduced here, with the associated quantities optical depth, scattering phase function and single scattering albedo. A distinction is made between so-called isotropic media, such as the atmosphere, and non-isotropic media such as vegetation canopies for which the radiative transfer equation is described in a different manner.

In chapter 3 radiative transfer models are classified into a number of categories on the basis of a number of criteria, such as the type of medium, the scattering behaviour and the manner in which the radiative transfer equation is represented. In addition, the most common solution methods, like ray tracing, successive orders of scattering approximation (SOSA), adding/doubling and analytical methods like the discrete ordinates method (DOM) are discussed.

A number of existing vegetation canopy reflectance models are reviewed in chapter 4, as well as a few atmospheric models. Here the criteria of chapter 3 are applied in order to classify these models into different categories.

A large part of this thesis is devoted to the so-called (N+2)-stream model, namely chapter 5, which introduces the theory behind this model, and chapter 6, which presents a large number of numerical and graphical results obtained for atmospheric scattering under a wide variety of conditions. In addition, in Appendices A and B many of the detailed mathematical formulations that form the basis of this model are presented.

The (N+2)-stream model has been developed by the author as a generalization and refinement of a previously existing four-stream model. The first example of a four-stream model is the vegetation canopy reflectance model of Suits (1972). The name four-stream model is derived from the fact that in such a model four radiant fluxes are considered, namely direct solar flux, direct flux (radiance) in the direction of the observer and two hemispherically diffuse fluxes (up- and downward). In the new model the direct fluxes are the same as in the four-stream model, but the two hemispherical fluxes have been replaced by N diffuse fluxes, hence the name (N+2)-stream model. This finer division of the sphere allows a more accurate calculation of the so-called multiple scattering contribution, especially when the scattering in the medium is highly anisotropic, such as in



a hazy atmosphere. What distinguishes this model from other existing radiative transfer formulations is that a special tessellation of the sphere of directions is proposed, the so-called equal-weight tessellation. Using this representation one obtains a uniform diffuse reflectance matrix for a Lambertian surface. Furthermore, in the analytical solution of the system of differential equations special attention is paid to the numerical problems that may be the result of duplicate eigenvalues. These problems have been satisfactorily solved now by the introduction of a limited number of suitable basis functions, and the numerical results presented in chapter 6 demonstrate that a robust numerical implementation is possible which performs well under a very broad range of input conditions. The mathematical validity of the analytical solution has been verified by the derivation of expressions for a layer's reflectance and transmittance which inherently obey Helmholtz's reciprocity relations. Also the behaviour of these expressions for small and infinite values of the optical thickness has been investigated to further confirm the validity of the model.

In chapter 7 a review of the SAIL vegetation canopy reflectance model is given. Like the model of Suits, the SAIL model is based on four-stream radiative transfer theory, but in SAIL the extinction and scattering coefficients are calculated for arbitrarily inclined leaves, whereas in the Suits model these are approximated by means of their horizontal and vertical projections. In this chapter also a new parameterization of the leaf inclination distribution function (LIDF) is introduced, which allows to characterize most existing distribution types found in reality by means of only two parameters. For a number of these distribution types the extinction and scattering coefficients are presented, as well as the associated bidirectional reflectance profiles as a function of the viewing angle. In Appendix C the theoretical background of the SAIL model is discussed with detailed derivations of the extinction and scattering coefficients.

Chapter 8 discusses the so-called hot spot effect and the extended model SAILH which includes this important phenomenon, based on the theory of A. Kuusk (1985). The hot spot effect is related to the finite size of the leaves in a canopy and it is explained by the fact that under viewing of the canopy along the direction opposite to the sunrays the fractions of observed shadowed leaf and soil area disappear. Under any other viewing direction always some shadowed leaf and soil area will be observed, so that at the hot spot a peak in the bidirectional reflectance occurs. By means of Kuusk's theory it is possible to predict how the reflectance decreases in directions



away from the hot spot, and how this behaviour is related to a parameter called hot spot size parameter, which is the ratio of horizontal correlation length and canopy height. The hot spot effect is illustrated by means of a number of examples showing bidirectional reflectance profiles in the principal plane of the sun. A theoretical limitation of Kuusk's approach is also discussed, as well as an attempt to improve the calculation of the effect. However, it appears that Kuusk's approach is sufficiently accurate under most circumstances, and the more accurate calculation increases the computation time considerably. The hot spot effect has interesting consequences for the reflectance (outside the hot spot) as a function of crop growth. It appears that crops following a different growth strategy (i.e. growing leaf area vs. making new leaves) also have different profiles of the reflectance as a function of the leaf area index (LAI). In particular for the first category (leaf growth) the prospects for LAI retrieval from remotely sensed data turn out to be much more favourable than was expected previously.

The application of four-stream radiative transfer theory to the terrestrial atmosphere for the purpose of correction of optical remote sensing imagery for atmospheric effects is discussed in chapter 9. The most important effects in the atmosphere are Rayleigh scattering by air molecules, Mie scattering by aerosol particles and selective absorption of radiation by water vapour, ozone and other atmospheric gases. These effects have been included in a model which computes the reflectance and transmittance properties of the atmospheric layer. The effect of the atmosphere on remotely sensed images can easily be described by means of the four-stream terminology and the adding method for the coupling of atmospheric parameters with the reflectance of the earth's surface. In this chapter a few methods to estimate atmospheric parameters from satellite images are discussed, and examples of this application to the atmospheric correction of Landsat Thematic Mapper images of the Netherlands are presented.

In chapter 10 an integrated soil-canopy-atmosphere radiative transfer model called OSCAR is introduced. This model is also based on four-stream theory and combines the atmosphere model of chapter 9 with the SAILH vegetation reflectance model. By means of this model one can simulate the normalized spectral radiance signal in any viewing direction and at any level within or above the atmosphere, so that also airborne remote sensing missions can be simulated. In addition, the so-called adjacency effect, which is caused by scattering of radiation from objects outside the field-of-view of the sensor into the observation beam, has been included here by allowing different



characteristics for target and background. Examples of simulations of this adjacency effect, and of the observation of different types of vegetation under various circumstances, are presented.

Chapter 11 gives a more general overview of applications of radiative transfer modelling to optical remote sensing problems. The main application fields fall in the categories prediction, explication and model inversion. Examples of modelling in these categories are discussed on the basis of data and simulation results in the red and near infrared parts of the spectrum. Also the application of vegetation indices is discussed, as well as applications in support of the mapping of the earth's albedo, which is important for climatological studies.

Appendix A Analytical solution of (N+2)-stream radiative transfer equation

The goal is to find the analytic solution of Eq. (5.13) for a homogeneous plane parallel layer in terms of reflectances and transmittances defined by Eq. (5.5) of chapter 5, or, given

$$\frac{d}{b dx} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o^+ \\ E_o^- \end{bmatrix} = \begin{bmatrix} k & & & & \\ -s' & A & -B & & \\ s & B & -A & & \\ w & v^T & v'^T & -K & \\ -w' & -v'^T & -v^T & & K \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \\ E_o^+ \\ E_o^- \end{bmatrix}, \quad (\text{A.1})$$

find the terms of the reflectance-transmittance matrix defined by

$$\begin{bmatrix} E_s(b) \\ E^-(b) \\ E^+(t) \\ E_o^+(t) \\ E_o^-(b) \end{bmatrix} = \begin{bmatrix} \tau_{ss} & & & & \\ \tau_{sd} & T & R & & \\ \rho_{sd} & R & T & & \\ \rho_{so} & \rho_{do}^T & \tau_{do}^T & \tau_{oo} & \\ \tau_{so} & \tau_{do}^T & \rho_{do}^T & & \tau_{oo} \end{bmatrix} \begin{bmatrix} E_s(t) \\ E^-(t) \\ E^+(b) \\ E_o^+(b) \\ E_o^-(t) \end{bmatrix}, \quad (\text{A.2})$$

where (b) and (t) refer to the bottom and top of the layer.

For the direct solar flux E_s one immediately finds $E_s(b) = E_s(t) e^{-kb}$, so $\tau_{ss} = e^{-kb}$.

For the diffuse flux vectors E^- and E^+ exclusively, the system of differential equations is reduced to

$$\frac{d}{b dx} \begin{bmatrix} E^- \\ E^+ \end{bmatrix} = \begin{bmatrix} A & -B \\ B & -A \end{bmatrix} \begin{bmatrix} E^- \\ E^+ \end{bmatrix}, \quad (\text{A.3})$$

in which A and B are matrices of dimension $\frac{1}{2}N \times \frac{1}{2}N$, and both E^- and E^+ have $\frac{1}{2}N$ elements.



As an example, a tessellation of the sphere of directions with $\Delta\varphi = \pi/6$ (azimuth sector width) and $\Delta\mu^2 = 1/3$ (zenith sector interval of cosine squared) is taken. This gives 12 azimuth sectors and 6 zenith zones, or $N = 72$. The elements of E^- and E^+ are ordered as follows:

$$\ell(E^-) = 12(i-1) + k, \text{ and}$$

$$\ell(E^+) = 12(6-i) + (k+5) \bmod 12 + 1,$$

where ℓ is the element number (1 to 36) in E^- or E^+ , and i and k are the zenith and azimuth sector numbers of the origin of the fluxes, where $i = 1, \dots, 6$ and $k = 1, \dots, 12$. See also Fig. 5.2 of section 5.2.

Corresponding elements of E^- and E^+ are now from opposite segments, which facilitates the solution by means of eigenvector decomposition. With this ordering Eq. (A.3) is written out as

$$\frac{d}{b dx} \begin{bmatrix} E_1^- \\ E_2^- \\ E_3^- \\ E_6^+ \\ E_5^+ \\ E_4^+ \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & -B_{16} & -B_{15} & -B_{14} \\ A_{21} & A_{22} & A_{23} & -B_{26} & -B_{25} & -B_{24} \\ A_{31} & A_{32} & A_{33} & -B_{36} & -B_{35} & -B_{34} \\ B_{61} & B_{62} & B_{63} & -A_{66} & -A_{65} & -A_{64} \\ B_{51} & B_{52} & B_{53} & -A_{56} & -A_{55} & -A_{54} \\ B_{41} & B_{42} & B_{43} & -A_{46} & -A_{45} & -A_{44} \end{bmatrix} \begin{bmatrix} E_1^- \\ E_2^- \\ E_3^- \\ E_6^+ \\ E_5^+ \\ E_4^+ \end{bmatrix}, \quad (\text{A.4})$$

in which the subscripts refer to the zenith zones. Each submatrix A_{ij} and B_{ij} describes the azimuthal scattering for incident zenith zone i and exitent zenith zone j .

The elements of A and B can be described by

$$A(i, j, \Delta k) = \kappa(i) \delta(i, j) \delta(\Delta k, 0) - \sigma(i, j, \Delta k), \text{ and}$$

$$B(i, j, \Delta k) = \sigma(i, j, \Delta k),$$

where i = exitent zenith zone

j = incident zenith zone

Δk = absolute difference in azimuth sector number

$\kappa(i)$ = diffuse extinction coefficient for zenith zone i



$\sigma(i, j, \Delta k)$ = scattering coefficient
 δ = Dirac delta function

Mirror symmetry with respect to the horizontal plane gives

$$\sigma(i, j, \Delta k) = \sigma(7-i, 7-j, \Delta k) \quad ,$$

whereas the reciprocity relation gives

$$\sigma(i, j, \Delta k) = \sigma(j, i, \Delta k) \quad .$$

Due to these symmetries the "A"-blocks and "B"-blocks of Eq. (A.4) are (apart from the sign) identical and in addition they are symmetric with respect to the diagonal.

The submatrices A_{ij} and B_{ij} are also symmetric, but since their elements depend only on Δk , their structure is even simpler. An example of such a matrix, called Z here, is given by

$$Z = \begin{bmatrix} a & b & c & d & e & f & g & f & e & d & c & b \\ b & a & b & c & d & e & f & g & f & e & d & c \\ c & b & a & b & c & d & e & f & g & f & e & d \\ d & c & b & a & b & c & d & e & f & g & f & e \\ e & d & c & b & a & b & c & d & e & f & g & f \\ f & e & d & c & b & a & b & c & d & e & f & g \\ g & f & e & d & c & b & a & b & c & d & e & f \\ f & g & f & e & d & c & b & a & b & c & d & e \\ e & f & g & f & e & d & c & b & a & b & c & d \\ d & e & f & g & f & e & d & c & b & a & b & c \\ c & d & e & f & g & f & e & d & c & b & a & b \\ b & c & d & e & f & g & f & e & d & c & b & a \end{bmatrix} .$$

The cyclic nature of this matrix suggests a decomposition by means of the discrete Fourier transform. For this, consider a complex vector w containing integer powers of the complex number $w = \cos n\varphi + i \sin n\varphi$, where $\varphi = \pi/6$, $i = (-1)^{1/2}$, and n is an integer greater than or equal to zero.



The elements of vector w are the series, $w^0, w^1, \dots, w^{11}$, which is a series of equally spaced points on the unit circle in the complex plane. Multiplication of this vector by matrix Z gives a vector consisting of the elements $z_1^T w$ to $z_{12}^T w$, where z_1^T to z_{12}^T are the rows of Z . For the first row this gives

$$z_1^T w = aw^0 + bw^1 + cw^2 + dw^3 + ew^4 + fw^5 + gw^6 + fw^7 + ew^8 + dw^9 + cw^{10} + bw^{11}.$$

The second row gives

$$z_2^T w = bw^0 + aw^1 + bw^2 + cw^3 + dw^4 + ew^5 + fw^6 + gw^7 + fw^8 + ew^9 + dw^{10} + cw^{11}.$$

However, since $w^0 = w^{12}$, this can be written as $w^1(z_1^T w)$, and in general, for row number j one obtains $z_j^T w = w^{j-1}(z_1^T w)$. The complete matrix-vector multiplication is written as

$$Zw = (z_1^T w)w = \lambda w.$$

Z transforms w into a scalar multiple of w , so w is an eigenvector of Z . The eigenvalue $\lambda = z_1^T w$ is a real number, since it can be written as

$$z_1^T w = aw^0 + gw^6 + b(w^1 + w^{11}) + c(w^2 + w^{10}) + d(w^3 + w^9) + e(w^4 + w^8) + f(w^5 + w^7).$$

Here, w^0 and w^6 are real numbers (equal to 1 and -1 , respectively), and the sums $w^j + w^{12-j}$ are also real, because w^j and w^{12-j} are complex conjugate pairs.

When n varies from zero to 11, one obtains 12 eigenvectors w , which can be placed in the eigenvector matrix W . This gives

$$ZW = W\Lambda, \tag{A.5}$$

where Λ is a diagonal matrix of eigenvalues.

Denoting the elements of W only by their power of $\cos(\pi/6) + i \sin(\pi/6)$, modulo 12, gives



$$W = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 2 & 4 & 6 & 8 & 10 & 0 & 2 & 4 & 6 & 8 & 10 \\ 0 & 3 & 6 & 9 & 0 & 3 & 6 & 9 & 0 & 3 & 6 & 9 \\ 0 & 4 & 8 & 0 & 4 & 8 & 0 & 4 & 8 & 0 & 4 & 8 \\ 0 & 5 & 10 & 3 & 8 & 1 & 6 & 11 & 4 & 9 & 2 & 7 \\ 0 & 6 & 0 & 6 & 0 & 6 & 0 & 6 & 0 & 6 & 0 & 0 \\ 0 & 7 & 2 & 9 & 4 & 11 & 6 & 1 & 8 & 3 & 10 & 5 \\ 0 & 8 & 4 & 0 & 8 & 4 & 0 & 8 & 4 & 0 & 8 & 4 \\ 0 & 9 & 6 & 3 & 0 & 9 & 6 & 3 & 0 & 9 & 6 & 3 \\ 0 & 10 & 8 & 6 & 4 & 2 & 0 & 10 & 8 & 6 & 4 & 2 \\ 0 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

From this it appears that W is also symmetric.

Now, consider the azimuthal scattering to be described by

$$\frac{d}{b dx} E_2 = Z E_1 \quad ,$$

where E_1 is an incident zonal flux vector of 12 elements, and E_2 is the scattered flux vector. For a transformed flux vector $E_2' = W E_2$ one obtains

$$\frac{d}{b dx} W E_2 = W Z E_1 = (Z W)^T E_1 = (W \Lambda)^T E_1 = \Lambda W E_1 \quad ,$$

where use was made of the symmetry of Z , W and Λ .

This means that for transformed zonal flux vectors $E_1' = W E_1$ and $E_2' = W E_2$ the azimuthal scattering can be described by means of a diagonal matrix Λ , and this is a great advantage. It may seem that using a complex transformation matrix is a disadvantage, but it turns out that a real matrix, consisting of the real part of W , is also sufficient. This is because in Eq. (A.5) Z and Λ are real, and taking real and imaginary parts leads to the same equations, but now for real and imaginary parts separately. For the real part



one obtains only the cosines, so the transformation used in that case is the discrete cosine transform, DCT. Because the cosine is an even function, rows 8 to 12 of the real part of matrix W are identical to rows 6 to 2, so rows 8 to 12 are redundant.

If the sun is located in the centre of azimuth sector one, then the diffuse flux field will be symmetric with respect to the principal plane, which means that the azimuth sector pairs (2,12), (3,11), (4,10), (5,9) and (6,8) have equal diffuse fluxes. This means that it is sufficient to consider only the fluxes in azimuth sectors 1 to 7 for a complete description of the diffuse flux field. In terms of the first seven azimuthal fluxes the DCT can be expressed by means of a matrix W defined by

$$W = \begin{pmatrix} 1 & 2 & 2 & 2 & 2 & 2 & 1 \\ 1 & \sqrt{3} & 1 & 0 & -1 & -\sqrt{3} & -1 \\ 1 & 1 & -1 & -2 & -1 & 1 & 1 \\ 1 & 0 & -2 & 0 & 2 & 0 & -1 \\ 1 & -1 & -1 & 2 & -1 & -1 & 1 \\ 1 & -\sqrt{3} & 1 & 0 & -1 & \sqrt{3} & -1 \\ 1 & -2 & 2 & -2 & 2 & -2 & 1 \end{pmatrix}.$$

This matrix is asymmetrical, but an advantage is that $W^2 = 12I$, where I is the 7×7 identity matrix, so that the inverse $W^{-1} = (1/12)W$.

By means of the above discrete cosine transform Eq. (A.4) is simplified to

$$\frac{d}{b dx} \begin{pmatrix} WE_1^- \\ WE_2^- \\ WE_3^- \\ WE_6^+ \\ WE_5^+ \\ WE_4^+ \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & -\Lambda_{16} & -\Lambda_{15} & -\Lambda_{14} \\ \Lambda_{12} & \Lambda_{22} & \Lambda_{23} & -\Lambda_{15} & -\Lambda_{25} & -\Lambda_{24} \\ \Lambda_{13} & \Lambda_{23} & \Lambda_{33} & -\Lambda_{14} & -\Lambda_{24} & -\Lambda_{34} \\ \Lambda_{16} & \Lambda_{15} & \Lambda_{14} & -\Lambda_{11} & -\Lambda_{12} & -\Lambda_{13} \\ \Lambda_{15} & \Lambda_{25} & \Lambda_{24} & -\Lambda_{12} & -\Lambda_{22} & -\Lambda_{23} \\ \Lambda_{14} & \Lambda_{24} & \Lambda_{34} & -\Lambda_{13} & -\Lambda_{23} & -\Lambda_{33} \end{pmatrix} \begin{pmatrix} WE_1^- \\ WE_2^- \\ WE_3^- \\ WE_6^+ \\ WE_5^+ \\ WE_4^+ \end{pmatrix}, \quad (A.6)$$

in which Λ_{ij} are 7×7 diagonal matrices containing the eigenvalues of the



azimuthal scattering matrices A_{ij} or B_{ij} . In the above numbering of subscripts the symmetries have been incorporated, so that it becomes more clear that only twelve distinct diagonal matrices are sufficient for a complete description of diffuse scattering. When the eigenvalues are represented by a vector λ of seven elements, it is given by

$$\lambda = Wz_1 ,$$

where z_1 contains the first 7 elements of the first column of the associate matrix A_{ij} or B_{ij} of Eq. (A.4).

By ordering Eq. (A.6) according to so-called azimuthal "modes" m , where $m = 0, \dots, 6$, one obtains for each mode

$$\frac{d}{b dx} \begin{bmatrix} E_m^- \\ E_m^+ \end{bmatrix} = \begin{bmatrix} A_m & -B_m \\ B_m & -A_m \end{bmatrix} \begin{bmatrix} E_m^- \\ E_m^+ \end{bmatrix} , \quad (A.7)$$

in which E_m^- and E_m^+ are azimuthally transformed zonal flux vectors for mode m . The matrices A_m and B_m describe only the zenith dependent scattering and have a dimension of only 3×3 .

This differential equation is identical in form to Eq. (A.3) but the dimension of the involved vectors and matrices has been reduced from 36 to 3 and this means a considerable reduction in storage space and computation times, even if seven of these equations have to be solved. This solution proceeds as follows:

Deleting the m subscript, one can write

$$\frac{d}{b dx} E^- = A E^- - B E^+ , \text{ and}$$

$$\frac{d}{b dx} E^+ = B E^- - A E^+ .$$

The sum and difference of both equations give



$$\frac{d}{b dx} (E^- + E^+) = (A + B) (E^- - E^+) \quad , \text{ and} \quad (\text{A.8.a})$$

$$\frac{d}{b dx} (E^- - E^+) = (A - B) (E^- + E^+) \quad . \quad (\text{A.8.b})$$

For $E^- + E^+$ a solution in the form of a linear combination of exponential functions is assumed, i.e.

$$E^- + E^+ = Y(e^{\Lambda bx} \delta_1 + e^{-\Lambda bx} \delta_2) \quad ,$$

where Y is a 3×3 matrix, δ_1 and δ_2 are constant vectors of dimension 3, and $e^{\Lambda bx}$ is a symbolic notation for the diagonal matrix $\text{diag} [\exp(\lambda_1 bx), \exp(\lambda_2 bx), \exp(\lambda_3 bx)]$.

Differentiating the above expression twice gives

$$\frac{d}{b dx} (E^- + E^+) = Y \Lambda (e^{\Lambda bx} \delta_1 - e^{-\Lambda bx} \delta_2) \quad , \text{ and}$$

$$\frac{d^2}{b^2 dx^2} (E^- + E^+) = Y \Lambda^2 (e^{\Lambda bx} \delta_1 + e^{-\Lambda bx} \delta_2) = Y \Lambda^2 Y^{-1} (E^- + E^+) \quad .$$

Differentiation of Eq. (A.8.a) gives

$$\frac{d^2}{b^2 dx^2} (E^- + E^+) = (A + B) \frac{d}{b dx} (E^- - E^+) = (A + B) (A - B) (E^- + E^+) \quad ,$$

so for a solution Y and Λ^2 should satisfy $(A + B) (A - B) = Y \Lambda^2 Y^{-1}$, or

$$(A + B) (A - B) Y = Y \Lambda^2 \quad ,$$

which means that matrix Y is a matrix of eigenvectors and Λ^2 contains the eigenvalues of $(A + B) (A - B)$, which in general is asymmetric. Therefore, this eigenvalue problem is solved as follows:

Matrix $(A + B)$, which is symmetric, is decomposed as $A + B = U \Delta U^T$, where Δ is a diagonal matrix of eigenvalues and U is a unitary matrix, i.e. $U U^T = I$.



$$\begin{aligned} \text{Now } U\Delta U^T(A-B) &= Y\Lambda^2 Y^{-1} , \\ \text{or } \Delta^{1/2} U^T(A-B) &= \Delta^{-1/2} U^T Y\Lambda^2 Y^{-1} , \quad (\text{where } \Delta^{1/2} \Delta^{1/2} = \Delta) \\ \text{or } \Delta^{1/2} U^T(A-B)U\Delta^{1/2} &= \Delta^{-1/2} U^T Y\Lambda^2 Y^{-1} U\Delta^{1/2} . \end{aligned}$$

Here, the matrix on the left is now symmetric, so that its eigenvector matrix V is unitary. Its eigenvalue matrix is Λ^2 , as can be seen from the expression on the right. So one may write

$$\Delta^{1/2} U^T(A-B)U\Delta^{1/2} = V\Lambda^2 V^T .$$

Equating corresponding matrices gives

$$\Delta^{-1/2} U^T Y = V , \text{ or } U^T Y = \Delta^{1/2} V , \text{ or } Y = U\Delta^{1/2} V .$$

The above decomposition is easy to implement by means of standard matrix diagonalization software (e.g. Jacobi's method) as only symmetric matrices are involved.

From $Y = U\Delta^{1/2} V$ it follows that $YY^T = U\Delta U^T = A+B$, and,

since $V^T \Delta^{1/2} U^T(A-B)U\Delta^{1/2} V = \Lambda^2 = Y^T(A-B)Y$, it follows that

$$Y^{-T} \Lambda^2 Y^{-1} = A-B , \text{ with } Y^{-T} = (Y^{-1})^T = (Y^T)^{-1} .$$

$$\begin{aligned} \text{From } \frac{d}{b dx} (E^- + E^+) &= Y\Lambda (e^{\Lambda bx} \delta_1 - e^{-\Lambda bx} \delta_2) = (A+B) (E^- - E^+) \\ &= YY^T (E^- - E^+) , \end{aligned}$$

$$\text{it follows that } E^- - E^+ = Y^{-T} \Lambda (e^{\Lambda bx} \delta_1 - e^{-\Lambda bx} \delta_2) ,$$

so that the complete solution is given by

$$E^- + E^+ = Y (e^{\Lambda bx} \delta_1 + e^{-\Lambda bx} \delta_2) , \text{ and}$$

$$E^- - E^+ = Y^{-T} \Lambda (e^{\Lambda bx} \delta_1 - e^{-\Lambda bx} \delta_2) .$$

Differentiation gives



$$\begin{aligned}
 \frac{d}{b dx} (E^- + E^+) &= Y \Lambda (e^{\Lambda b x} \delta_1 - e^{-\Lambda b x} \delta_2) = Y \Lambda \Lambda^{-1} Y^T (E^- - E^+) = Y Y^T (E^- - E^+) \\
 &= (A + B) (E^- - E^+) \quad , \text{ and} \\
 \frac{d}{b dx} (E^- - E^+) &= Y^{-T} \Lambda^2 (e^{\Lambda b x} \delta_1 + e^{-\Lambda b x} \delta_2) = Y^{-T} \Lambda^2 Y^{-1} (E^- + E^+) \\
 &= (A - B) (E^- + E^+) \quad ,
 \end{aligned}$$

so these solutions are formally correct. However, for mode $m = 0$ the matrix $A - B$ becomes singular in the case of conservative scattering, i.e. in that case $A - B$ has an eigenvalue equal to zero. This follows from the equation $F \mathbf{1} + B \mathbf{1} = \omega \kappa \mathbf{1}$, given in chapter 5. Since $A = \kappa - F$, this equation leads to

$$(\kappa - A) \mathbf{1} + B \mathbf{1} = \omega \kappa \mathbf{1} \quad , \text{ or } (A - B) \mathbf{1} = (1 - \omega) \kappa \mathbf{1} .$$

For conservative scattering the single scattering albedo ω equals one, so that $(A - B) \mathbf{1} = 0$, so $\mathbf{1}$ is an eigenvector of $A - B$, with eigenvalue zero. In this case

$$\frac{d}{b dx} \mathbf{1}^T (E^- - E^+) = \mathbf{1}^T (A - B) (E^- + E^+) = 0 \quad , \text{ so that } \mathbf{1}^T (E^- - E^+)$$

is constant.

Also , $\frac{d}{b dx} \mathbf{1}^T (A + B)^{-1} (E^- + E^+) = \mathbf{1}^T (E^- - E^+) = \text{constant}$, so that the component $\mathbf{1}^T (A + B)^{-1} (E^- + E^+)$ becomes linear in x .

However, this does not follow from the general solution in terms of exponential functions. Therefore, an alternative solution, which covers also the case of conservative scattering for $m = 0$, is introduced.

This solution is based on hyperbolic sine and cosine functions which are taken relative to the centre of the layer ($x = -1/2$), here defined by

$$\begin{aligned}
 \mathbf{C}(x) &= e^{\Lambda b x} + e^{-\Lambda b(1+x)} \quad , \text{ and} \\
 \mathbf{S}(x) &= \Lambda^{-1} [e^{\Lambda b x} - e^{-\Lambda b(1+x)}] \quad .
 \end{aligned}$$

When one of the eigenvalue λ equals zero, the associate element of $\mathbf{C}(x)$

becomes equal to 2 , and that of $S(x)$ to $(1+2x)b$, a linear function in x .
Differentiation of the above functions gives

$$\frac{d}{b dx} C(x) = \Lambda [e^{\Lambda b x} - e^{-\Lambda b(1+x)}] = \Lambda^2 S(x) \quad , \text{ and}$$

$$\frac{d}{b dx} S(x) = \Lambda^{-1} \Lambda [e^{\Lambda b x} + e^{-\Lambda b(1+x)}] = C(x) \quad .$$

The general solution is now formulated as

$$E^- + E^+ = Y [C(x)\delta_1 + S(x)\delta_2] \quad , \quad (\text{A.9.a})$$

$$E^- - E^+ = Y^{-T} [\Lambda^2 S(x)\delta_1 + C(x)\delta_2] \quad . \quad (\text{A.9.b})$$

Differentiation gives

$$\begin{aligned} \frac{d}{b dx} (E^- + E^+) &= Y [\Lambda^2 S(x)\delta_1 + C(x)\delta_2] = Y Y^T (E^- - E^+) \\ &= (A + B) (E^- - E^+) \quad , \end{aligned}$$

$$\begin{aligned} \frac{d}{b dx} (E^- - E^+) &= Y^{-T} [\Lambda^2 C(x)\delta_1 + \Lambda^2 S\delta_2] = Y^{-T} \Lambda^2 Y^{-1} (E^- + E^+) \\ &= (A - B) (E^- + E^+) \quad . \end{aligned}$$

Therefore, Eqs. (A.9) are also formally correct. The case of conservative scattering for azimuthal mode $m = 0$ gives $(A - B) \mathbf{1} = \mathbf{0}$, so also

$$(A + B) (A - B) \mathbf{1} = \mathbf{1} \cdot \mathbf{0}$$

Since, in general, $(A + B) (A - B) Y = Y \Lambda^2$, it is concluded that the eigenvector from Y associated with the zero eigenvalue must be a scalar multiple of the one-vector $\mathbf{1}$. Since Eq. (A.9.b) gives

$$\begin{aligned} Y^T (E^- - E^+) &= \Lambda^2 S(x)\delta_1 + C(x)\delta_2 \quad , \text{ the component with eigenvalue zero} \\ \text{gives } \mathbf{1}^T (E^- - E^+) &= 2\delta_{21} \quad , \text{ which is indeed constant, as expected.} \end{aligned}$$

Here δ_{21} is the element of vector δ_2 which corresponds to the column of Y that contains the eigenvector with eigenvalue zero.



Also, Eq. (A.9.a) gives

$$\begin{aligned}(A+B)^{-1}(E^-+E^+) &= (YY^T)^{-1}Y[C(x)\delta_1+S(x)\delta_2] \\ &= Y^{-T}[C(x)\delta_1+S(x)\delta_2] \quad , \text{ or} \\ Y^T(A+B)^{-1}(E^-+E^+) &= C(x)\delta_1+S(x)\delta_2 \quad .\end{aligned}$$

Here, the component with eigenvalue zero is equal to

$$\mathbf{1}^T(A+B)^{-1}(E^-+E^+) = 2\delta_{11} + (1+2x)b\delta_{21} \quad , \text{ which is indeed linear in } x \quad .$$

The reason that the case of conservative scattering leads to a zero eigenvalue only for $m = 0$ is that only for this azimuthal mode all azimuthal fluxes within a zenith zone are added together, so that the vector $\mathbf{1}^T$ applied to the zonal fluxes gives the sum of all fluxes over a hemisphere.

The next step is to introduce the direct solar flux, E_s , which leads to the generalized Duntley-equations, i.e.

$$\frac{d}{b dx} \begin{bmatrix} E_s \\ E^- \\ E^+ \end{bmatrix} = \begin{bmatrix} k & & \\ -s' & A & -B \\ s & B & -A \end{bmatrix} \begin{bmatrix} E_s \\ E^- \\ E^+ \end{bmatrix} \quad .$$

Application of the discrete cosine transform to the azimuthal fluxes per zenith zone gives

$$\begin{aligned}\frac{d}{b dx} E_m^- &= -s'_m E_s + A_m E_m^- - B_m E_m^+ \quad , \\ \frac{d}{b dx} E_m^+ &= s_m E_s + B_m E_m^- - A_m E_m^+ \quad .\end{aligned}$$

Omitting the m subscript and taking the sum and difference of these equations gives



$$\frac{d}{b dx} (E^- + E^+) = -(s' - s) E_s + (A + B) (E^- - E^+) \quad ,$$

$$\frac{d}{b dx} (E^- - E^+) = -(s' + s) E_s + (A - B) (E^- + E^+) \quad .$$

Transformation by means of Y^{-1} and Y^T yields

$$\begin{aligned} \frac{d}{b dx} Y^{-1} (E^- + E^+) &= -Y^{-1} (s' - s) E_s + Y^{-1} (A + B) (E^- - E^+) \\ &= -Y^{-1} (s' - s) E_s + Y^T (E^- - E^+) \quad , \text{ and} \end{aligned} \quad (\text{A.10.a})$$

$$\begin{aligned} \frac{d}{b dx} Y^T (E^- - E^+) &= -Y^T (s' + s) E_s + Y^T (A - B) (E^- + E^+) \\ &= -Y^T (s' + s) E_s + \Lambda^2 Y^{-1} (E^- + E^+) \quad . \end{aligned} \quad (\text{A.10.b})$$

Differentiation of the former equation and substitution of the latter gives

$$\frac{d^2}{b^2 dx^2} Y^{-1} (E^- + E^+) = -k Y^{-1} (s' - s) E_s - Y^T (s' + s) E_s + \Lambda^2 Y^{-1} (E^- + E^+) \quad .$$

A general solution is given by

$$Y^{-1} (E^- + E^+) = e^{\Lambda b x} \delta_1 + e^{-\Lambda b x} \delta_2 + c E_s \quad ,$$

where c is a constant vector of the correct dimension.

Differentiating the general solution twice gives

$$\frac{d^2}{b^2 dx^2} Y^{-1} (E^- + E^+) = \Lambda^2 Y^{-1} (E^- + E^+) + (k^2 - \Lambda^2) c E_s \quad ,$$

in which k is the diagonal matrix formed by $k = k I$. Comparison of both previous differential equations gives for the vector c

$$c = -(k^2 - \Lambda^2)^{-1} [k Y^{-1} (s' - s) + Y^T (s' + s)] \quad .$$

From Eq. (A.10.a) one can derive

$$\begin{aligned} Y^T(E^- - E^+) &= \frac{d}{b dx} Y^{-1}(E^- + E^+) + Y^{-1}(s' - s)E_s \\ &= \Lambda e^{\Lambda b x} \delta_1 - \Lambda e^{-\Lambda b x} \delta_2 + [k c + Y^{-1}(s' - s)] E_s \quad . \end{aligned}$$

Putting $d = k c + Y^{-1}(s' - s)$ summarizes the general solution as

$$Y^{-1}(E^- + E^+) = e^{\Lambda b x} \delta_1 + e^{-\Lambda b x} \delta_2 + c E_s \quad , \quad (\text{A.11.a})$$

$$Y^T(E^- - E^+) = \Lambda e^{\Lambda b x} \delta_1 - \Lambda e^{-\Lambda b x} \delta_2 + d E_s \quad . \quad (\text{A.11.b})$$

Substitution of c gives for d

$$d = -(k^2 - \Lambda^2)^{-1} [\Lambda^2 Y^{-1}(s' - s) + k Y^T] \quad .$$

Elements of c and d can become infinite if an eigenvalue λ equals k , the extinction coefficient for direct solar flux. This potential numerical problem is avoided as follows:

Writing $p_s = (k + \Lambda)^{-1} Y^{-1}(s' - s)$ and $q_s = (k + \Lambda)^{-1} Y^T(s' + s)$, c and d can be expressed as

$$\begin{aligned} c &= -(k - \Lambda)^{-1} [k p_s + q_s] \quad \text{and} \quad d = -(k - \Lambda)^{-1} [\Lambda^2 p_s + k q_s] \quad , \text{ or} \\ c &= -p_s - (k - \Lambda)^{-1} [\Lambda p_s + q_s] \quad \text{and} \quad d = -q_s - \Lambda (k - \Lambda)^{-1} [\Lambda p_s + q_s] \quad . \end{aligned}$$

With $E_s = E_s(0) e^{k b x}$, where $E_s(0)$ is the solar irradiance at the top, this gives for Eqs. (A.11)



$$\begin{aligned}
 Y^{-1}(E^{-}+E^{+}) &= e^{\Lambda bx} \delta_1 + e^{-\Lambda bx} \delta_2 - (k-\Lambda)^{-1} [\Lambda p_s + q_s] E_s(0) e^{kbx} \\
 &\quad - p_s E_s(0) e^{kbx} , \\
 Y^T(E^{-}-E^{+}) &= \Lambda e^{\Lambda bx} \delta_1 - \Lambda e^{-\Lambda bx} \delta_2 - \Lambda (k-\Lambda)^{-1} [\Lambda p_s + q_s] E_s(0) e^{kbx} \\
 &\quad - q_s E_s(0) e^{kbx} .
 \end{aligned}$$

In both equations the term $e^{\Lambda bx} \delta_1 - (k-\Lambda)^{-1} [\Lambda p_s + q_s] E_s(0) e^{kbx}$ can be rearranged to

$$\begin{aligned}
 &e^{\Lambda bx} \delta_1 - (k-\Lambda)^{-1} [e^{kbx} - e^{\Lambda bx} + e^{\Lambda bx}] [\Lambda p_s + q_s] E_s(0) \\
 &= e^{\Lambda bx} [\delta_1 - (k-\Lambda)^{-1} (\Lambda p_s + q_s) E_s(0)] \\
 &\quad + (k-\Lambda)^{-1} [e^{\Lambda bx} - e^{kbx}] [\Lambda p_s + q_s] E_s(0) \\
 &= e^{\Lambda bx} \delta_1' + J(x) [\Lambda p_s + q_s] E_s(0) ,
 \end{aligned}$$

$$\text{with } \delta_1' = \delta_1 - (k-\Lambda)^{-1} (\Lambda p_s + q_s) E_s(0) ,$$

$$\text{and } J(x) = (k-\Lambda)^{-1} [e^{\Lambda bx} - e^{kbx}] .$$

By this, the infinities caused by $k = \lambda$ have been removed, as the expression $(k-\lambda)^{-1} [e^{\lambda bx} - e^{kbx}]$ for $k = \lambda$ can be replaced by the limit for $k \rightarrow \lambda$, which equals $-bx e^{kbx}$, and vector δ_1' can be considered finite in all circumstances. In order to avoid exponentials with large positive arguments, the vector δ_2 is replaced by $\delta_2' = e^{\Lambda b} \delta_2$, which gives

$$\begin{aligned}
 Y^{-1}(E^{-}+E^{+}) &= e^{\Lambda bx} \delta_1' + e^{-\Lambda b(1+x)} \delta_2' + J(x) [\Lambda p_s + q_s] E_s(0) \\
 &\quad - p_s E_s(0) e^{kbx} , \text{ and} \\
 Y^T(E^{-}-E^{+}) &= \Lambda e^{\Lambda bx} \delta_1' - \Lambda e^{-\Lambda b(1+x)} \delta_2' + \Lambda J(x) [\Lambda p_s + q_s] E_s(0) \\
 &\quad - q_s E_s(0) e^{kbx} .
 \end{aligned}$$

As x is negative inside the layer, all exponentials have zero or negative arguments in this case.

Finally, in order to avoid problems with the case of conservative scattering for $m = 0$, which gives $\lambda = 0$, the functions $C(x)$ and $S(x)$ are



reintroduced by putting $\delta_1'' = \frac{1}{2}(\delta_1' + \delta_2')$ and $\delta_2'' = \frac{1}{2}\Lambda(\delta_1' - \delta_2')$.
Omitting the double primes, the final result is given by

$$Y^{-1}(E^{-} + E^{+}) = C(x)\delta_1 + S(x)\delta_2 + J(x) [\Lambda p_s + q_s] E_s(0) - p_s E_s(0) e^{kbx}, \text{ and} \quad (\text{A.12.a})$$

$$Y^T(E^{-} - E^{+}) = \Lambda^2 S(x)\delta_1 + C(x)\delta_2 + \Lambda J(x) [\Lambda p_s + q_s] E_s(0) - q_s E_s(0) e^{kbx}. \quad (\text{A.12.b})$$

Differentiation of $J(x)$ gives

$$\begin{aligned} \frac{d}{b dx} J(x) &= (k - \Lambda)^{-1} [\Lambda e^{\Lambda bx} - k e^{kbx}] \\ &= (k - \Lambda)^{-1} [\Lambda (e^{\Lambda bx} - e^{kbx}) - (k - \Lambda) e^{kbx}] \\ &= \Lambda J(x) - e^{kbx}. \end{aligned}$$

Therefore, differentiation of Eq. (A.12.a) gives

$$\begin{aligned} \frac{d}{b dx} Y^{-1}(E^{-} + E^{+}) &= \Lambda^2 S(x)\delta_1 + C(x)\delta_2 + \Lambda J(x) [\Lambda p_s + q_s] E_s(0) \\ &\quad - e^{kbx} [\Lambda p_s + q_s] E_s(0) - k p_s E_s(0) e^{kbx} \\ &= Y^T(E^{-} - E^{+}) - (k + \Lambda) p_s E_s(0) e^{kbx} \\ &= Y^T(E^{-} - E^{+}) - Y^{-1}(s' - s) E_s, \text{ or} \\ \frac{d}{b dx} (E^{-} + E^{+}) &= -(s' - s) E_s + Y Y^T(E^{-} + E^{+}) \\ &= -(s' - s) E_s + (A + B) (E^{-} - E^{+}). \end{aligned}$$

Similarly, differentiation of Eq. (A.12.b) gives



$$\begin{aligned}
\frac{d}{b dx} Y^T (E^- - E^+) &= \Lambda^2 C(x) \delta_1 + \Lambda^2 S(x) \delta_2 + \Lambda^2 J(x) [\Lambda p_s + q_s] E_s(0) \\
&\quad - e^{k b x} \Lambda [\Lambda p_s + q_s] E_s(0) - k q_s E_s(0) e^{k b x} \\
&= \Lambda^2 Y^{-1} (E^- + E^+) - (k + \Lambda) q_s E_s(0) e^{k b x} \\
&= \Lambda^2 Y^{-1} (E^- + E^+) - Y^T (s' + s) E_s \quad , \text{ or} \\
\frac{d}{b dx} (E^- - E^+) &= -(s' + s) E_s + Y^{-T} \Lambda^2 Y^{-1} (E^- + E^+) \\
&= -(s' + s) E_s + (A - B) (E^- + E^+) \quad .
\end{aligned}$$

This proves that the solutions satisfy the differential equations.

Eqs. (A.12) will now be evaluated at the top ($x = 0$) and the bottom ($x = -1$) in order to derive the reflectances and transmittances. First, for the functions $C(x)$, $S(x)$ and $J(x)$ one finds

$$\begin{aligned}
C(0) &= I + e^{-\Lambda b} & ; & \quad C(-1) = e^{-\Lambda b} + I = C(0) & ; \\
S(0) &= \Lambda^{-1} [I - e^{-\Lambda b}] & ; & \quad S(-1) = \Lambda^{-1} [e^{-\Lambda b} - I] = -S(0) & ; \\
J(0) &= O & ; & \quad J(-1) = (k - \Lambda)^{-1} [e^{-\Lambda b} - e^{-k b}] \quad .
\end{aligned}$$

where I is the identity matrix and O the zero matrix.

In what follows the notations $C = C(0)$, $S = S(0)$, and $J_1(k) = J(-1)$ will be used. Now the boundary equations are given by

$$Y^{-1} [E^-(0) + E^+(0)] = C \delta_1 + S \delta_2 - p_s E_s(0) \quad , \quad (\text{A.13.a})$$

$$\begin{aligned}
Y^{-1} [E^-(-1) + E^+(-1)] &= C \delta_1 - S \delta_2 + J_1(k) [\Lambda p_s + q_s] E_s(0) \\
&\quad - p_s E_s(0) e^{-k b} \quad , \quad (\text{A.13.b})
\end{aligned}$$

$$Y^T [E^-(0) - E^+(0)] = \Lambda^2 S \delta_1 + C \delta_2 - q_s E_s(0) \quad , \quad (\text{A.13.c})$$



$$Y^T [E^-(1) - E^+(1)] = -\Lambda^2 S \delta_1 + C \delta_2 + \Lambda J_1(k) [\Lambda p_s + q_s] E_s(0) - q_s E_s(0) e^{-kb} \quad . \quad (A.13.d)$$

From the sum and difference of Eqs. (A.13.a-b) one obtains

$$Y^{-1} [E^-(0) + E^+(-1) + E^-(-1) + E^+(0)] = 2C \delta_1 - g_1^+ E_s(0) \quad ,$$

$$Y^{-1} [E^-(0) - E^+(-1) - E^-(-1) + E^+(0)] = 2S \delta_2 - g_1^- E_s(0) \quad ,$$

$$\text{where } g_1^+ = (1 + e^{-kb}) p_s - J_1(k) (\Lambda p_s + q_s) \quad ,$$

$$\text{and } g_1^- = (1 - e^{-kb}) p_s + J_1(k) (\Lambda p_s + q_s) \quad .$$

Similarly, the sum and difference of Eqs. (A.13.c-d) yield

$$Y^T [E^-(0) - E^+(-1) + E^-(-1) - E^+(0)] = 2C \delta_2 - g_2^- E_s(0) \quad ,$$

$$Y^T [E^-(0) + E^+(-1) - E^-(-1) - E^+(0)] = 2\Lambda^2 S \delta_1 - g_2^+ E_s(0) \quad ,$$

$$\text{where } g_2^- = (1 + e^{-kb}) q_s - \Lambda J_1(k) (\Lambda p_s + q_s) \quad ,$$

$$\text{and } g_2^+ = (1 - e^{-kb}) q_s + \Lambda J_1(k) (\Lambda p_s + q_s) \quad .$$

Now, premultiplication by suitable matrices allows to rewrite the boundary equations as

$$S \Lambda^2 Y^{-1} [E^-(0) + E^+(-1) + E^-(-1) + E^+(0)] = 2S \Lambda^2 C \delta_1 - S \Lambda^2 g_1^+ E_s(0) \quad ,$$

$$C Y^T [E^-(0) + E^+(-1) - E^-(-1) - E^+(0)] = 2C \Lambda^2 S \delta_1 - C g_2^+ E_s(0) \quad ,$$

$$C Y^{-1} [E^-(0) - E^+(-1) - E^-(-1) + E^+(0)] = 2C S \delta_2 - C g_1^- E_s(0) \quad ,$$

$$S Y^T [E^-(0) - E^+(-1) + E^-(-1) - E^+(0)] = 2S C \delta_2 - S g_2^- E_s(0) \quad .$$

The vectors δ_1 and δ_2 can now be eliminated since S , Λ^2 and C are diagonal matrices, so that $S \Lambda^2 C = C \Lambda^2 S$ and $CS = SC$.

The difference of the first pair of equations then gives



$$[\mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1} - \mathbf{C}\mathbf{Y}^T][\mathbf{E}^-(0) + \mathbf{E}^+(-1)] + [\mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1} + \mathbf{C}\mathbf{Y}^T][\mathbf{E}^-(-1) + \mathbf{E}^+(0)] \\ = [\mathbf{C}\mathbf{g}_2^+ - \mathbf{S}\mathbf{A}^2\mathbf{g}_1^+]E_s(0) \text{ , or}$$

$$\mathbf{E}^-(-1) + \mathbf{E}^+(0) = \\ [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1}]^{-1} \{ [\mathbf{C}\mathbf{Y}^T - \mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1}][\mathbf{E}^-(0) + \mathbf{E}^+(-1)] + [\mathbf{C}\mathbf{g}_2^+ - \mathbf{S}\mathbf{A}^2\mathbf{g}_1^+]E_s(0) \} \text{ .}$$

The difference of the second pair of equations gives

$$[\mathbf{S}\mathbf{Y}^T - \mathbf{C}\mathbf{Y}^{-1}][\mathbf{E}^-(0) - \mathbf{E}^+(-1)] + [\mathbf{S}\mathbf{Y}^T + \mathbf{C}\mathbf{Y}^{-1}][\mathbf{E}^-(-1) - \mathbf{E}^+(0)] \\ = [\mathbf{C}\mathbf{g}_1^- - \mathbf{S}\mathbf{g}_2^-]E_s(0) \text{ , or}$$

$$\mathbf{E}^-(-1) - \mathbf{E}^+(0) = \\ [\mathbf{C}\mathbf{Y}^{-1} + \mathbf{S}\mathbf{Y}^T]^{-1} \{ [\mathbf{C}\mathbf{Y}^{-1} - \mathbf{S}\mathbf{Y}^T][\mathbf{E}^-(0) - \mathbf{E}^+(-1)] + [\mathbf{C}\mathbf{g}_1^- - \mathbf{S}\mathbf{g}_2^-]E_s(0) \} \text{ .}$$

Returning to Eq. (A.2) one finds

$$\mathbf{E}^-(-1) = \tau_{sd}E_s(0) + \mathbf{T}\mathbf{E}^-(0) + \mathbf{R}\mathbf{E}^+(-1) \text{ , and} \\ \mathbf{E}^+(0) = \rho_{sd}E_s(0) + \mathbf{R}\mathbf{E}^-(0) + \mathbf{T}\mathbf{E}^+(-1) \text{ .}$$

The sum and difference give

$$\mathbf{E}^-(-1) + \mathbf{E}^+(0) = (\tau_{sd} + \rho_{sd})E_s(0) + (\mathbf{T} + \mathbf{R})[\mathbf{E}^-(0) + \mathbf{E}^+(-1)] \text{ , and} \\ \mathbf{E}^-(-1) - \mathbf{E}^+(0) = (\tau_{sd} - \rho_{sd})E_s(0) + (\mathbf{T} - \mathbf{R})[\mathbf{E}^-(0) - \mathbf{E}^+(-1)] \text{ .}$$

Comparing these with the previous results leads to the conclusion that

$$\mathbf{T} + \mathbf{R} = [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1}]^{-1} [\mathbf{C}\mathbf{Y}^T - \mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1}] \text{ ,} \quad (\text{A.14})$$

$$\mathbf{T} - \mathbf{R} = [\mathbf{C}\mathbf{Y}^{-1} + \mathbf{S}\mathbf{Y}^T]^{-1} [\mathbf{C}\mathbf{Y}^{-1} - \mathbf{S}\mathbf{Y}^T] \text{ ,} \quad (\text{A.15})$$

$$\tau_{sd} + \rho_{sd} = [\mathbf{C}\mathbf{Y}^T + \mathbf{S}\mathbf{A}^2\mathbf{Y}^{-1}]^{-1} [\mathbf{C}\mathbf{g}_2^+ - \mathbf{S}\mathbf{A}^2\mathbf{g}_1^+] \text{ ,} \quad (\text{A.16})$$

$$\tau_{sd} - \rho_{sd} = [\mathbf{C}\mathbf{Y}^{-1} + \mathbf{S}\mathbf{Y}^T]^{-1} [\mathbf{C}\mathbf{g}_1^- - \mathbf{S}\mathbf{g}_2^-] \text{ .} \quad (\text{A.17})$$

Alternative expressions for the above quantities are obtained by first solving the vectors δ_1 and δ_2 from the boundary equations. For this, the boundary



equations are written as

$$\begin{aligned}
 E^-(0) + E^+(-1) + E^-(-1) + E^+(0) &= 2YC\delta_1 - Yg_1^+ E_s(0) \quad , \\
 E^-(0) + E^+(-1) - E^-(-1) - E^+(0) &= 2Y^{-T}\Lambda^2 S\delta_1 - Y^{-T}g_2^+ E_s(0) \quad , \\
 E^-(0) - E^+(-1) - E^-(-1) + E^+(0) &= 2YS\delta_2 - Yg_1^- E_s(0) \quad , \\
 E^-(0) - E^+(-1) + E^-(-1) - E^+(0) &= 2Y^{-T}C\delta_2 - Y^{-T}g_2^- E_s(0) \quad .
 \end{aligned}$$

The sum of the first pair of equations gives

$$\delta_1 = [YC + Y^{-T}\Lambda^2 S]^{-1} \{ E^-(0) + E^+(-1) + \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] E_s(0) \} \quad , \quad (A.18)$$

and the sum of the second pair gives

$$\delta_2 = [YS + Y^{-T}C]^{-1} \{ E^-(0) - E^+(-1) + \frac{1}{2} [Yg_1^- + Y^{-T}g_2^-] E_s(0) \} \quad , \quad (A.19)$$

The difference of the first pair gives

$$\begin{aligned}
 E^-(-1) + E^+(0) &= [YC - Y^{-T}\Lambda^2 S] \delta_1 - \frac{1}{2} [Yg_1^+ - Y^{-T}g_2^+] E_s(0) \\
 &= [YC - Y^{-T}\Lambda^2 S] [YC + Y^{-T}\Lambda^2 S]^{-1} \{ E^-(0) + E^+(-1) + \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] E_s(0) \} \\
 &\quad - \frac{1}{2} [Yg_1^+ - Y^{-T}g_2^+] E_s(0) \quad .
 \end{aligned}$$

Since from Eq. (A.2) it follows that

$$E^-(-1) + E^+(0) = (T+R) [E^-(0) + E^+(-1)] + (\tau_{sd} + \rho_{sd}) E_s(0) \quad , \text{ one finds}$$

$$T+R = [YC - Y^{-T}\Lambda^2 S] [YC + Y^{-T}\Lambda^2 S]^{-1} \quad , \text{ and} \quad (A.20)$$

$$\tau_{sd} + \rho_{sd} = (T+R) \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] - \frac{1}{2} [Yg_1^+ - Y^{-T}g_2^+] \quad . \quad (A.21)$$

Similarly, the difference of the second pair gives

$$\begin{aligned}
 E^-(-1) - E^+(0) &= [Y^{-T}C - YS] \delta_2 + \frac{1}{2} [Yg_1^- - Y^{-T}g_2^-] E_s(0) \\
 &= [Y^{-T}C - YS] [Y^{-T}C + YS]^{-1} \{ E^-(0) - E^+(-1) + \frac{1}{2} [Yg_1^- + Y^{-T}g_2^-] E_s(0) \} \\
 &\quad - \frac{1}{2} [Yg_1^- - Y^{-T}g_2^-] E_s(0) \\
 &= (T-R) [E^-(0) - E^+(-1)] + (\tau_{sd} + \rho_{sd}) E_s(0) \quad , \text{ so that}
 \end{aligned}$$



$$\mathbf{T} - \mathbf{R} = [\mathbf{Y}^{-\text{T}}\mathbf{C} - \mathbf{YS}] [\mathbf{Y}^{-\text{T}}\mathbf{C} + \mathbf{YS}]^{-1} \quad , \text{ and} \quad (\text{A.22})$$

$$\tau_{sd} - \rho_{sd} = (\mathbf{T} - \mathbf{R})^{1/2} [\mathbf{Y}\mathbf{g}_1^- + \mathbf{Y}^{-\text{T}}\mathbf{g}_2^-] + 1/2 [\mathbf{Y}\mathbf{g}_1^- - \mathbf{Y}^{-\text{T}}\mathbf{g}_2^-] \quad . \quad (\text{A.23})$$

The equivalence of the expressions for $\mathbf{T} + \mathbf{R}$, Eqs. (A.14) and (A.20), and those for $\mathbf{T} - \mathbf{R}$, Eqs. (A.15) and (A.22), is only possible if $\mathbf{T}$ and $\mathbf{R}$ are symmetric matrices, in which case transposing Eq. (A.14) leads to Eq. (A.20) and transposing Eq. (A.15) leads to Eq. (A.22). It is therefore concluded that $\mathbf{T}$ and $\mathbf{R}$ are symmetric.

Eqs. (A.21) and (A.23) look quite different from Eqs. (A.16) and (A.17), respectively. However, substitution of $\mathbf{T} + \mathbf{R}$ of Eq. (A.14) in Eq. (A.21), and of $\mathbf{T} - \mathbf{R}$ of Eq. (A.15) in Eq. (A.23) gives

$$\begin{aligned} \tau_{sd} + \rho_{sd} &= [\mathbf{CY}^{\text{T}} + \mathbf{SA}^2\mathbf{Y}^{-1}]^{-1} \{ [\mathbf{CY}^{\text{T}} - \mathbf{SA}^2\mathbf{Y}^{-1}]^{1/2} [\mathbf{Y}\mathbf{g}_1^+ + \mathbf{Y}^{-\text{T}}\mathbf{g}_2^+] \\ &\quad - [\mathbf{CY}^{\text{T}} + \mathbf{SA}^2\mathbf{Y}^{-1}]^{1/2} [\mathbf{Y}\mathbf{g}_1^+ - \mathbf{Y}^{-\text{T}}\mathbf{g}_2^+] \} \\ &= [\mathbf{CY}^{\text{T}} + \mathbf{SA}^2\mathbf{Y}^{-1}]^{-1} [\mathbf{Cg}_2^+ - \mathbf{SA}^2\mathbf{g}_1^+] \quad , \text{ and} \end{aligned}$$

$$\begin{aligned} \tau_{sd} - \rho_{sd} &= [\mathbf{CY}^{-1} + \mathbf{SY}^{\text{T}}]^{-1} \{ [\mathbf{CY}^{-1} - \mathbf{SY}^{\text{T}}]^{1/2} [\mathbf{Y}\mathbf{g}_1^- + \mathbf{Y}^{-\text{T}}\mathbf{g}_2^-] \\ &\quad + [\mathbf{CY}^{-1} + \mathbf{SY}^{\text{T}}]^{1/2} [\mathbf{Y}\mathbf{g}_1^- - \mathbf{Y}^{-\text{T}}\mathbf{g}_2^-] \} \\ &= [\mathbf{CY}^{-1} + \mathbf{SY}^{\text{T}}]^{-1} [\mathbf{Cg}_2^- - \mathbf{Sg}_1^-] \quad , \end{aligned}$$

by which Eqs. (A.16) and (A.17) have been reobtained.

This concludes the solution of the generalized Duntley-equations for a single layer, which gives the diffuse reflectances and transmittances for diffuse incidence ($\mathbf{R}$ and $\mathbf{T}$) and specular incidence (ρ_{sd} and τ_{sd}).

The directional reflectances and transmittances for diffuse incident fluxes and those for specular incidence (called bidirectional reflectance and transmittance) are solved from the differential equations



$$\begin{aligned}\frac{d}{b dx} E_o^+ &= w E_s + \mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+ - K E_o^+ \quad , \text{ and} \\ \frac{d}{b dx} E_o^- &= -w' E_s - \mathbf{v}'^T \mathbf{E}^- - \mathbf{v}^T \mathbf{E}^+ + K E_o^- \quad .\end{aligned}$$

In what follows it is assumed that E_o^+ and E_o^- are in exactly opposite directions. According to Eq.(A.2), the result will be expressed by

$$\begin{aligned}E_o^+(0) &= \rho_{so} E_s(0) + \rho_{do}^T \mathbf{E}^-(0) + \tau_{do}^T \mathbf{E}^+(-1) + \tau_{oo} E_o^+(-1) \quad , \text{ and} \\ E_o^-(-1) &= \tau_{so} E_s(0) + \tau_{do}^T \mathbf{E}^-(0) + \rho_{do}^T \mathbf{E}^+(-1) + \tau_{oo} E_o^-(0) \quad .\end{aligned}$$

For the solution, first the fluxes E_o^+ and E_o^- are separated into single scattering contributions E_o^{+s} and E_o^{-s} , and multiple scattering contributions E_o^{+d} and E_o^{-d} .

For the single scattering contributions the diffuse flux vectors $\mathbf{E}^-$ and $\mathbf{E}^+$ are ignored, giving

$$\begin{aligned}\frac{d}{b dx} E_o^{+s} &= w E_s - K E_o^{+s} \quad , \text{ and} \\ \frac{d}{b dx} E_o^{-s} &= -w' E_s + K E_o^{-s} \quad , \text{ of which the solutions are} \\ E_o^{+s}(x) &= e^{-Kbx} \left[\int_{-1}^x w E_s e^{Kbx} dx + E_o^{+s}(-1) e^{-Kb} \right] \quad , \text{ and} \\ E_o^{-s}(x) &= e^{Kbx} \left[\int_x^0 w' E_s e^{-Kbx} dx + E_o^{-s}(0) \right] \quad .\end{aligned}$$

with $E_s = E_s(0)e^{kbx}$, this gives

$$\begin{aligned}E_o^{+s}(x) &= \frac{w e^{-Kbx}}{K+k} \left[e^{(K+k)bx} - e^{-(K+k)b} \right] E_s(0) + E_o^{+s}(-1) e^{-Kb(1+x)} \quad , \text{ and} \\ E_o^{-s}(x) &= \frac{w'}{K-k} \left[e^{kbx} - e^{Kbx} \right] E_s(0) + E_o^{-s}(0) e^{Kbx} \quad .\end{aligned}$$

In the latter expression the function $Q^-(x)$, defined by

$Q^-(x) = (e^{kbx} - e^{Kbx})/(K-k)$, becomes indeterminate for $K = k$, so in that case the limit for $K \rightarrow k$ must be taken, which is equal to $-bx e^{kbx}$.



For E_o^{+s} at the top ($x = 0$) and E_o^{-s} at the bottom of the layer ($x = -1$) one obtains

$$E_o^{+s}(0) = \frac{w}{K+k} [1 - e^{-(K+k)b}] E_s(0) + E_o^{+s}(-1)e^{-Kb} \quad , \text{ and}$$

$$E_o^{-s}(-1) = \frac{w'}{K-k} [e^{-kb} - e^{-Kb}] E_s(0) + E_o^{-s}(0)e^{-Kb} \quad ,$$

which can be equated to

$$E_o^{+s}(0) = \rho_{so}^s E_s(0) + \tau_{oo} E_o^{+s}(-1) \quad , \text{ and}$$

$$E_o^{-s}(-1) = \tau_{so}^s E_s(0) + \tau_{oo} E_o^{-s}(0) \quad .$$

Here τ_{oo} is the direct transmittance in the direction of observation and ρ_{so}^s and τ_{so}^s are the single scattering contributions to ρ_{so} and τ_{so} , respectively. Comparison of the corresponding equations gives

$$\tau_{oo} = e^{-Kb} \quad ,$$

$$\rho_{so}^s = \frac{w}{K+k} [1 - e^{-(K+k)b}] \quad ,$$

$$\tau_{so}^s = \frac{w'}{K-k} [e^{-kb} - e^{-Kb}] \quad (\text{for } K \neq k) \quad ,$$

$$\tau_{so}^s = w' b e^{-kb} \quad (\text{for } K = k) \quad .$$

For the multiple scattering contributions the specular flux is ignored in the differential equations, leading to

$$\frac{d}{b dx} E_o^{+d} = \mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+ - K E_o^{+d} \quad , \text{ and}$$

$$\frac{d}{b dx} E_o^{-d} = -\mathbf{v}'^T \mathbf{E}^- - \mathbf{v}^T \mathbf{E}^+ + K E_o^{-d} \quad , \text{ of which the solutions are}$$

$$E_o^{+d}(x) = e^{-Kbx} \left[\int_{-1}^x (\mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+) e^{Kbx} dx + E_o^{+d}(-1)e^{-Kb} \right] \quad , \text{ and}$$

$$E_o^{-d}(x) = e^{Kbx} \left[\int_x^{-1} (\mathbf{v}'^T \mathbf{E}^- + \mathbf{v}^T \mathbf{E}^+) e^{-Kbx} dx + E_o^{-d}(0) \right] \quad .$$

For E_o^{+d} at the top ($x = 0$) and E_o^{-d} at the bottom of the layer ($x = -1$) this



gives

$$E_o^{+d}(0) = \int_{-1}^0 (\mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+) e^{Kbx} dx + E_o^{+d}(-1) e^{-Kb} \quad , \text{ and}$$

$$E_o^{-d}(-1) = e^{-Kb} \int_{-1}^0 (\mathbf{v}'^T \mathbf{E}^- + \mathbf{v}^T \mathbf{E}^+) e^{-Kbx} dx + E_o^{-d}(0) e^{-Kb} \quad .$$

This can be equated to

$$E_o^{+d}(0) = \rho_{so}^d E_s(0) + \rho_{do}^T \mathbf{E}^-(0) + \tau_{do}^T \mathbf{E}^+(-1) + \tau_{oo} E_o^{+d}(-1) \quad , \text{ and}$$

$$E_o^{-d}(-1) = \tau_{so}^d E_s(0) + \tau_{do}^T \mathbf{E}^-(0) + \rho_{do}^T \mathbf{E}^+(-1) + \tau_{oo} E_o^{-d}(0) \quad .$$

From this it follows again that $\tau_{oo} = e^{-Kb}$ and

$$\rho_{so}^d E_s(0) + \rho_{do}^T \mathbf{E}^-(0) + \tau_{do}^T \mathbf{E}^+(-1) = \int_{-1}^0 (\mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+) e^{Kbx} dx = G_1 \quad , \text{ and}$$

$$\tau_{so}^d E_s(0) + \tau_{do}^T \mathbf{E}^-(0) + \rho_{do}^T \mathbf{E}^+(-1) = e^{-Kb} \int_{-1}^0 (\mathbf{v}'^T \mathbf{E}^- + \mathbf{v}^T \mathbf{E}^+) e^{-Kbx} dx = G_2 \quad .$$

Sum and difference of the above equations give

$$(\rho_{so}^d + \tau_{so}^d) E_s(0) + (\rho_{do}^T + \tau_{do}^T) [\mathbf{E}^-(0) + \mathbf{E}^+(-1)] = G_1 + G_2 \quad , \text{ and} \quad (\text{A.24.a})$$

$$(\rho_{so}^d - \tau_{so}^d) E_s(0) + (\rho_{do}^T - \tau_{do}^T) [\mathbf{E}^-(0) - \mathbf{E}^+(-1)] = G_1 - G_2 \quad . \quad (\text{A.24.b})$$

Calculation of the integrals G_1 and G_2 is facilitated by writing

$$\mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+ = \frac{1}{2} (\mathbf{v}^T + \mathbf{v}'^T) (\mathbf{E}^- + \mathbf{E}^+) + \frac{1}{2} (\mathbf{v}^T - \mathbf{v}'^T) (\mathbf{E}^- - \mathbf{E}^+) \quad , \text{ and}$$

$$\mathbf{v}'^T \mathbf{E}^- + \mathbf{v}^T \mathbf{E}^+ = \frac{1}{2} (\mathbf{v}^T + \mathbf{v}'^T) (\mathbf{E}^- + \mathbf{E}^+) - \frac{1}{2} (\mathbf{v}^T - \mathbf{v}'^T) (\mathbf{E}^- - \mathbf{E}^+) \quad .$$

By analogy with the vectors $\mathbf{p}_s$ and $\mathbf{q}_s$, associated with s and s' by

$$\mathbf{p}_s = (\mathbf{k} + \mathbf{\Lambda})^{-1} \mathbf{Y}^{-1} (s' - s) \quad \text{and} \quad \mathbf{q}_s = (\mathbf{k} + \mathbf{\Lambda})^{-1} \mathbf{Y}^T (s' + s) \quad ,$$

one can define the (transposed) vectors $\mathbf{p}_v^T$ and $\mathbf{q}_v^T$, associated with $\mathbf{v}^T$ and $\mathbf{v}'^T$ by



$$\mathbf{p}_v^T = (\mathbf{v}'^T - \mathbf{v}^T) \mathbf{Y}^{-T} (\mathbf{K} + \mathbf{\Lambda})^{-1} \quad \text{and} \quad \mathbf{q}_v^T = (\mathbf{v}'^T + \mathbf{v}^T) \mathbf{Y} (\mathbf{K} + \mathbf{\Lambda})^{-1} \quad , \text{ which gives}$$

$$\mathbf{v}^T + \mathbf{v}'^T = \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^{-1} \quad \text{and} \quad \mathbf{v}^T - \mathbf{v}'^T = -\mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^T \quad .$$

Hence,

$$\mathbf{v}^T \mathbf{E}^- + \mathbf{v}'^T \mathbf{E}^+ = \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^{-1} (\mathbf{E}^- + \mathbf{E}^+) - \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^T (\mathbf{E}^- - \mathbf{E}^+) \quad , \text{ and}$$

$$\mathbf{v}'^T \mathbf{E}^- + \mathbf{v}^T \mathbf{E}^+ = \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^{-1} (\mathbf{E}^- + \mathbf{E}^+) + \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{Y}^T (\mathbf{E}^- - \mathbf{E}^+) \quad .$$

By defining the following integral vectors

$$\mathbf{f}_\rho^+ = \int_{-1}^0 \mathbf{Y}^{-1} (\mathbf{E}^- + \mathbf{E}^+) \mathbf{e}^{Kbx} dx \quad , \quad \mathbf{f}_\tau^+ = \mathbf{e}^{-Kb} \int_{-1}^0 \mathbf{Y}^{-1} (\mathbf{E}^- + \mathbf{E}^+) \mathbf{e}^{-Kbx} dx \quad ,$$

$$\mathbf{f}_\rho^- = \int_{-1}^0 \mathbf{Y}^T (\mathbf{E}^- - \mathbf{E}^+) \mathbf{e}^{Kbx} dx \quad , \quad \mathbf{f}_\tau^- = \mathbf{e}^{-Kb} \int_{-1}^0 \mathbf{Y}^T (\mathbf{E}^- - \mathbf{E}^+) \mathbf{e}^{-Kbx} dx \quad ,$$

one can determine G_1 and G_2 as

$$G_1 = \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{f}_\rho^+ - \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{f}_\rho^- \quad , \text{ and}$$

$$G_2 = \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{f}_\tau^+ + \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) \mathbf{f}_\tau^- \quad .$$

For the calculation of the vectors $\mathbf{f}_\rho^+$, $\mathbf{f}_\rho^-$, $\mathbf{f}_\tau^+$ and $\mathbf{f}_\tau^-$ Eqs. (A.12) must be integrated. The result is based on the integrals of the functions $\mathbf{e}^{kbx}$, $\mathbf{C}(x)$, $\mathbf{S}(x)$ and $\mathbf{J}(x)$, which can be summarized by



$$\int_{-1}^0 e^{kbx} e^{Kbx} dx = (K+k)^{-1} [1 - e^{-(K+k)b}] = Q^+ ;$$

$$e^{-Kb} \int_{-1}^0 e^{kbx} e^{-Kbx} dx = (K-k)^{-1} [e^{-kb} - e^{-Kb}] = Q^- ;$$

$$\int_{-1}^0 C(x) e^{Kbx} dx = J_1(K) + J_2(K) ;$$

$$e^{-Kb} \int_{-1}^0 C(x) e^{-Kbx} dx = J_1(K) + J_2(K) ;$$

$$\int_{-1}^0 S(x) e^{Kbx} dx = -\Lambda^{-1} [J_1(K) - J_2(K)] ;$$

$$e^{-Kb} \int_{-1}^0 S(x) e^{-Kbx} dx = \Lambda^{-1} [J_1(K) - J_2(K)] ;$$

$$\int_{-1}^0 J(x) e^{Kbx} dx = -(K+\Lambda)^{-1} [J_1(k) e^{-Kb} - Q^+] = D^+ ;$$

$$e^{-Kb} \int_{-1}^0 J(x) e^{-Kbx} dx = (K-k)^{-1} [J_1(k) - J_1(K)] = D^- .$$

In the above expressions, $Q^+ = Q^+ I$, and $J_1(k)$, $J_1(K)$ and $J_2(K)$ are defined by

$$J_1(k) = (k - \Lambda)^{-1} [e^{-\Lambda b} - e^{-kb}] ,$$

$$J_1(K) = (K - \Lambda)^{-1} [e^{-\Lambda b} - e^{-Kb}] , \text{ and}$$

$$J_2(K) = (K + \Lambda)^{-1} [I - e^{-(K+\Lambda)b}] .$$

All quantities are diagonal matrices, except Q^+ and Q^- , which are scalars, but these can also be treated as diagonal matrices when the context requires it.



Application of the above results gives

$$\begin{aligned}
 f_\rho^+ &= L^+ \delta_1 - L^- \delta_2 + [D^+(\Lambda p_s + q_s) - Q^+ p_s] E_s(0) \quad ; \\
 f_\tau^+ &= L^+ \delta_1 + L^- \delta_2 + [D^-(\Lambda p_s + q_s) - Q^- p_s] E_s(0) \quad ; \\
 f_\rho^- &= -\Lambda^2 L^- \delta_1 + L^+ \delta_2 + [\Lambda D^+(\Lambda p_s + q_s) - Q^+ q_s] E_s(0) \quad ; \\
 f_\tau^- &= \Lambda^2 L^- \delta_1 + L^+ \delta_2 + [\Lambda D^-(\Lambda p_s + q_s) - Q^- q_s] E_s(0) \quad ;
 \end{aligned}$$

$$\text{where } L^+ = J_1(K) + J_2(K) \quad \text{and} \quad L^- = \Lambda^{-1} [J_1(K) - J_2(K)] \quad .$$

Sums and differences of the above equations give

$$\begin{aligned}
 f_\rho^+ + f_\tau^+ &= 2L^+ \delta_1 + [(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-) p_s] E_s(0) \quad ; \\
 f_\rho^+ - f_\tau^+ &= -2L^- \delta_2 + [(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-) p_s] E_s(0) \quad ; \\
 f_\rho^- + f_\tau^- &= 2L^+ \delta_2 + [\Lambda(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-) q_s] E_s(0) \quad ; \\
 f_\rho^- - f_\tau^- &= -2\Lambda^2 L^- \delta_1 + [\Lambda(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-) q_s] E_s(0) \quad .
 \end{aligned}$$

The sum and difference of G_1 and G_2 are given by

$$\begin{aligned}
 G_1 + G_2 &= \frac{1}{2} q_v^T (K + \Lambda) (f_\rho^+ + f_\tau^+) - \frac{1}{2} p_v^T (K + \Lambda) (f_\rho^- - f_\tau^-) \quad , \text{ and} \\
 G_1 - G_2 &= \frac{1}{2} q_v^T (K + \Lambda) (f_\rho^+ - f_\tau^+) - \frac{1}{2} p_v^T (K + \Lambda) (f_\rho^- + f_\tau^-) \quad ,
 \end{aligned}$$

and substitution of the f -vectors yields

$$\begin{aligned}
 G_1 + G_2 &= [q_v^T (K + \Lambda) L^+ + p_v^T (K + \Lambda) \Lambda^2 L^-] \delta_1 \\
 &\quad + \frac{1}{2} q_v^T (K + \Lambda) [(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-) p_s] E_s(0) \\
 &\quad - \frac{1}{2} p_v^T (K + \Lambda) [\Lambda(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-) q_s] E_s(0) \quad , \text{ and}
 \end{aligned}$$

$$\begin{aligned}
 G_1 - G_2 &= [-q_v^T (K + \Lambda) L^- - p_v^T (K + \Lambda) L^+] \delta_2 \\
 &\quad + \frac{1}{2} q_v^T (K + \Lambda) [(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-) p_s] E_s(0) \\
 &\quad - \frac{1}{2} p_v^T (K + \Lambda) [\Lambda(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-) q_s] E_s(0) \quad .
 \end{aligned}$$

The vectors δ_1 and δ_2 have been given by Eqs. (A.18) and (A.19), respectively, and are repeated here:



$$\begin{aligned}\delta_1 &= [YC + Y^{-T}\Lambda^2S]^{-1} \{ E^-(0) + E^+(-1) + \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] E_s(0) \} ; \\ \delta_2 &= [YS + Y^{-T}C]^{-1} \{ E^-(0) - E^+(-1) + \frac{1}{2} [Yg_1^- + Y^{-T}g_2^-] E_s(0) \} .\end{aligned}$$

Now, Eqs. (A.24) can be written out as follows:

$$\begin{aligned}(\rho_{so}^d + \tau_{so}^d) E_s(0) + (\rho_{do}^T + \tau_{do}^T) [E^-(0) + E^+(-1)] = \\ [q_v^T(K + \Lambda)L^+ + p_v^T(K + \Lambda)\Lambda^2L^-] * \\ [YC + Y^{-T}\Lambda^2S]^{-1} \{ E^-(0) + E^+(-1) + \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] E_s(0) \} \\ + \frac{1}{2} q_v^T(K + \Lambda) [(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-)p_s] E_s(0) \\ - \frac{1}{2} p_v^T(K + \Lambda) [\Lambda(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-)q_s] E_s(0) , \text{ and}\end{aligned}$$

$$\begin{aligned}(\rho_{so}^d - \tau_{so}^d) E_s(0) + (\rho_{do}^T - \tau_{do}^T) [E^-(0) - E^+(-1)] = \\ -[q_v^T(K + \Lambda)L^- + p_v^T(K + \Lambda)L^+] * \\ [YS + Y^{-T}C]^{-1} \{ E^-(0) - E^+(-1) + \frac{1}{2} [Yg_1^- + Y^{-T}g_2^-] E_s(0) \} \\ + \frac{1}{2} q_v^T(K + \Lambda) [(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-)p_s] E_s(0) \\ - \frac{1}{2} p_v^T(K + \Lambda) [\Lambda(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-)q_s] E_s(0) .\end{aligned}$$

Equating corresponding quantities gives

$$\rho_{do}^T + \tau_{do}^T = [q_v^T(K + \Lambda)L^+ + p_v^T(K + \Lambda)\Lambda^2L^-] [YC + Y^{-T}\Lambda^2S]^{-1} , \quad (A.25)$$

$$\begin{aligned}\rho_{so}^d + \tau_{so}^d = (\rho_{do}^T + \tau_{do}^T) \frac{1}{2} [Yg_1^+ + Y^{-T}g_2^+] \\ + \frac{1}{2} q_v^T(K + \Lambda) [(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-)p_s] \\ - \frac{1}{2} p_v^T(K + \Lambda) [\Lambda(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-)q_s] , \quad (A.26)\end{aligned}$$

$$\rho_{do}^T - \tau_{do}^T = -[q_v^T(K + \Lambda)L^- + p_v^T(K + \Lambda)L^+] [YS + Y^{-T}C]^{-1} , \quad (A.27)$$

$$\begin{aligned} \rho_{so}^d - \tau_{so}^d &= (\rho_{do}^T - \tau_{do}^T) \frac{1}{2} [Yg_1^- + Y^{-T}g_2^-] \\ &\quad + \frac{1}{2} q_v^T (K + \Lambda) [(D^+ - D^-)(\Lambda p_s + q_s) - (Q^+ - Q^-)p_s] \\ &\quad - \frac{1}{2} p_v^T (K + \Lambda) [\Lambda(D^+ + D^-)(\Lambda p_s + q_s) - (Q^+ + Q^-)q_s] \quad . \quad (A.28) \end{aligned}$$

The four expressions above can be used to calculate ρ_{so}^d , τ_{so}^d , ρ_{do}^T and τ_{do}^T . However, the reciprocity relations require that the couples (ρ_{do}, ρ_{sd}) and (τ_{do}, τ_{sd}) commute when the directions of specular incidence and of observation are interchanged. Also, ρ_{so}^d and τ_{so}^d should be invariant with respect to this. From the given expressions these properties are far from obvious. In order to prove the reciprocity relations, alternative expressions are developed now.

First, by analogy with the g -vectors, which are expressed in p_s , q_s and k , the (transposed) h -vectors are introduced as

$$\begin{aligned} h_1^{+T} &= p_v^T (1 + e^{-Kb}) - (p_v^T \Lambda + q_v^T) J_1(K) \quad ; \\ h_2^{+T} &= q_v^T (1 - e^{-Kb}) + (p_v^T \Lambda + q_v^T) \Lambda J_1(K) \quad ; \\ h_1^{-T} &= p_v^T (1 - e^{-Kb}) + (p_v^T \Lambda + q_v^T) J_1(K) \quad ; \\ h_2^{-T} &= q_v^T (1 + e^{-Kb}) - (p_v^T \Lambda + q_v^T) \Lambda J_1(K) \quad . \end{aligned}$$

Writing $z_1^T = q_v^T (K + \Lambda) L^+ + p_v^T (K + \Lambda) \Lambda^2 L^-$, and $-z_2^T = q_v^T (K + \Lambda) L^- + p_v^T (K + \Lambda) L^+$ gives

$$\begin{aligned} \rho_{do}^T + \tau_{do}^T &= z_1^T [YC + Y^{-T} \Lambda^2 S]^{-1} \quad , \text{ and} \\ \rho_{do}^T - \tau_{do}^T &= z_2^T [YS + Y^{-T} C]^{-1} \quad . \end{aligned}$$

The (transposed) vectors z_1^T and z_2^T can be rearranged as follows:



$$\begin{aligned}
z_1^T &= q_v^T (K + \Lambda) L^+ + p_v^T (K + \Lambda) \Lambda^2 L^- \\
&= \{ q_v^T [J_1(K) + J_2(K)] + p_v^T \Lambda [J_1(K) - J_2(K)] \} (K + \Lambda) \\
&= q_v^T [(K - \Lambda) J_1(K) + (K + \Lambda) J_2(K)] \\
&\quad - p_v^T \Lambda [-(K - \Lambda) J_1(K) + (K + \Lambda) J_2(K)] \\
&\quad + (p_v^T \Lambda + q_v^T) J_1(K) 2\Lambda \\
&= q_v^T [e^{-\Lambda b} - e^{-Kb} + I - e^{-(K+\Lambda)b}] - p_v^T \Lambda [-e^{-\Lambda b} + e^{-Kb} + I - e^{-(K+\Lambda)b}] \\
&\quad + (p_v^T \Lambda + q_v^T) J_1(K) 2\Lambda \\
&= q_v^T (I - e^{-Kb}) C - p_v^T \Lambda (I + e^{-Kb}) \Lambda S + (p_v^T \Lambda + q_v^T) J_1(K) [C + \Lambda S] \Lambda \\
&= [q_v^T (1 - e^{-Kb}) + (p_v^T \Lambda + q_v^T) \Lambda J_1(K)] C \\
&\quad - [p_v^T (1 + e^{-Kb}) - (p_v^T \Lambda + q_v^T) J_1(K)] \Lambda^2 S \\
&= h_2^{+T} C - h_1^{+T} \Lambda^2 S
\end{aligned}$$

$$\begin{aligned}
-z_2^T &= q_v^T (K + \Lambda) L^- + p_v^T (K + \Lambda) L^+ \\
&= \{ q_v^T \Lambda^{-1} [J_1(K) - J_2(K)] + p_v^T [J_1(K) + J_2(K)] \} (K + \Lambda) \\
&= q_v^T \Lambda^{-1} [(K - \Lambda) J_1(K) - (K + \Lambda) J_2(K)] \\
&\quad + p_v^T [(K - \Lambda) J_1(K) + (K + \Lambda) J_2(K)] \\
&\quad + (q_v^T \Lambda^{-1} + p_v^T) J_1(K) 2\Lambda \\
&= q_v^T \Lambda^{-1} [e^{-\Lambda b} - e^{-Kb} - I + e^{-(K+\Lambda)b}] + p_v^T [e^{-\Lambda b} - e^{-Kb} + I - e^{-(K+\Lambda)b}] \\
&\quad + (p_v^T \Lambda + q_v^T) J_1(K) 2 \\
&= -q_v^T \Lambda^{-1} (I + e^{-Kb}) \Lambda S + p_v^T (I - e^{-Kb}) C + (p_v^T \Lambda + q_v^T) J_1(K) [C + \Lambda S] \\
&= [p_v^T (1 - e^{-Kb}) + (p_v^T \Lambda + q_v^T) J_1(K)] C \\
&\quad - [q_v^T (1 + e^{-Kb}) - (p_v^T \Lambda + q_v^T) \Lambda J_1(K)] S \\
&= h_1^{-T} C - h_2^{-T} S
\end{aligned}$$

The result of these rearrangements is expressed by

$$\rho_{do}^T + \tau_{do}^T = [h_2^{+T} C - h_1^{+T} \Lambda^2 S] [YC + Y^{-T} \Lambda^2 S]^{-1}, \text{ and} \quad (\text{A.29})$$

$$\rho_{do}^T - \tau_{do}^T = [h_2^{-T} S - h_1^{-T} C] [YS + Y^{-T} C]^{-1}. \quad (\text{A.30})$$

Transposing the above expressions and replacing the h -vectors with the corresponding g -vectors gives results identical with Eqs. (A.16) and (A.17),



which proves the reciprocity relation for the couples (ρ_{do}, ρ_{sd}) and (τ_{do}, τ_{sd}) .

In order to prove the reciprocity relations for ρ_{so}^d and τ_{so}^d , first Eqs. (A.29) and (A.30) are rearranged and subsequently substituted in Eqs. (A.26) and (A.28). For this, Eqs. (A.20) and (A.22) are applied, which give

$$\begin{aligned} T+R &= [YC - Y^{-T}\Lambda^2S][YC + Y^{-T}\Lambda^2S]^{-1} \\ &= I - 2Y^{-T}\Lambda^2S[YC + Y^{-T}\Lambda^2S]^{-1} \\ &= -I + 2YC[YC + Y^{-T}\Lambda^2S]^{-1}, \text{ and} \end{aligned}$$

$$\begin{aligned} T-R &= [Y^{-T}C - YS][Y^{-T}C + YS]^{-1} \\ &= I - 2YS[Y^{-T}C + YS]^{-1} \\ &= -I + 2Y^{-T}C[Y^{-T}C + YS]^{-1}. \end{aligned}$$

Now, Eqs. (A.29) and (A.30) are written as

$$\begin{aligned} \rho_{do}^T + \tau_{do}^T &= [h_2^{+T}Y^{-1}YC - h_1^{+T}Y^TY^{-T}\Lambda^2S][YC + Y^{-T}\Lambda^2S]^{-1} \\ &= \frac{1}{2}h_2^{+T}Y^{-1}(T+R+I) + \frac{1}{2}h_1^{+T}Y^T(T+R-I) \\ &= \frac{1}{2}[h_2^{+T}Y^{-1} + h_1^{+T}Y^T](R+T) + \frac{1}{2}[h_2^{+T}Y^{-1} - h_1^{+T}Y^T], \text{ and} \quad (\text{A.31}) \end{aligned}$$

$$\begin{aligned} \rho_{do}^T - \tau_{do}^T &= [h_2^{-T}Y^{-1}YS - h_1^{-T}Y^TY^{-T}C][Y^{-T}C + YS]^{-1} \\ &= \frac{1}{2}h_2^{-T}Y^{-1}(I-T+R) + \frac{1}{2}h_1^{-T}Y^T(-I-T+R) \\ &= \frac{1}{2}[h_2^{-T}Y^{-1} + h_1^{-T}Y^T](R-T) + \frac{1}{2}[h_2^{-T}Y^{-1} - h_1^{-T}Y^T]. \quad (\text{A.32}) \end{aligned}$$

These equations are similar to Eqs. (A.21) and (A.23), the alternative expressions for $\tau_{sd} + \rho_{sd}$ and $\tau_{sd} - \rho_{sd}$, respectively.

Substitution of Eqs. (A.31) and (A.32) in Eqs. (A.26) and (A.28), respectively, gives



$$\begin{aligned}
\rho_{so}^d + \tau_{so}^d = & \frac{1}{2} [\mathbf{h}_2^{+T} \mathbf{Y}^{-1} + \mathbf{h}_1^{+T} \mathbf{Y}^T] (\mathbf{R} + \mathbf{T}) \frac{1}{2} [\mathbf{Y} \mathbf{g}_1^+ + \mathbf{Y}^{-T} \mathbf{g}_2^+] \\
& + \frac{1}{2} [\mathbf{h}_2^{+T} \mathbf{Y}^{-1} - \mathbf{h}_1^{+T} \mathbf{Y}^T] \frac{1}{2} [\mathbf{Y} \mathbf{g}_1^+ + \mathbf{Y}^{-T} \mathbf{g}_2^+] \\
& + \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) [(\mathbf{D}^+ + \mathbf{D}^-) (\mathbf{\Lambda} \mathbf{p}_s + \mathbf{q}_s) - (\mathbf{Q}^+ + \mathbf{Q}^-) \mathbf{p}_s] \\
& - \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) [\mathbf{\Lambda} (\mathbf{D}^+ - \mathbf{D}^-) (\mathbf{\Lambda} \mathbf{p}_s + \mathbf{q}_s) - (\mathbf{Q}^+ - \mathbf{Q}^-) \mathbf{q}_s] , \quad (\text{A.33})
\end{aligned}$$

$$\begin{aligned}
\rho_{so}^d - \tau_{so}^d = & \frac{1}{2} [\mathbf{h}_2^{-T} \mathbf{Y}^{-1} + \mathbf{h}_1^{-T} \mathbf{Y}^T] (\mathbf{R} - \mathbf{T}) \frac{1}{2} [\mathbf{Y} \mathbf{g}_1^- + \mathbf{Y}^{-T} \mathbf{g}_2^-] \\
& + \frac{1}{2} [\mathbf{h}_2^{-T} \mathbf{Y}^{-1} - \mathbf{h}_1^{-T} \mathbf{Y}^T] \frac{1}{2} [\mathbf{Y} \mathbf{g}_1^- + \mathbf{Y}^{-T} \mathbf{g}_2^-] \\
& + \frac{1}{2} \mathbf{q}_v^T (\mathbf{K} + \mathbf{\Lambda}) [(\mathbf{D}^+ - \mathbf{D}^-) (\mathbf{\Lambda} \mathbf{p}_s + \mathbf{q}_s) - (\mathbf{Q}^+ - \mathbf{Q}^-) \mathbf{p}_s] \\
& - \frac{1}{2} \mathbf{p}_v^T (\mathbf{K} + \mathbf{\Lambda}) [\mathbf{\Lambda} (\mathbf{D}^+ + \mathbf{D}^-) (\mathbf{\Lambda} \mathbf{p}_s + \mathbf{q}_s) - (\mathbf{Q}^+ + \mathbf{Q}^-) \mathbf{q}_s] . \quad (\text{A.34})
\end{aligned}$$

As both equations consist of four terms, they are written as

$$\rho_{so}^d + \tau_{so}^d = t_1^+ + t_2^+ + t_3^+ + t_4^+ , \text{ and}$$

$$\rho_{so}^d - \tau_{so}^d = t_1^- + t_2^- + t_3^- + t_4^- .$$

The first terms, t_1^+ and t_1^- , are completely symmetric with respect to the directions of specular incidence and observation, and therefore will be left alone. The second terms, t_2^+ and t_2^- , are of the form

$$t_2 = \frac{1}{4} [\mathbf{h}_2^T \mathbf{Y}^{-1} \mathbf{Y}^{-T} \mathbf{g}_2 - \mathbf{h}_1^T \mathbf{Y}^T \mathbf{Y} \mathbf{g}_1] + \frac{1}{4} [\mathbf{h}_2^T \mathbf{g}_1 - \mathbf{h}_1^T \mathbf{g}_2] ,$$

where the '+' and '-' superscripts have been omitted. The first part of this expression on the right is symmetric, whereas the second part is anti-symmetric, i.e. it changes sign when the directions are interchanged. In both equations, this anti-symmetric part should now be combined with the terms t_3 and t_4 , and lead to a symmetric result. Elaboration of this for Eq. (A.33) gives



$$\begin{aligned}
& \frac{1}{4} [h_2^{+T} g_1^+ - h_1^{+T} g_2^+] + t_3^+ + t_4^+ = -\frac{1}{4} (p_v^T \Lambda + q_v^T) (K-k) D^-(\Lambda p_s + q_s) \\
& - \frac{1}{4} (p_v^T \Lambda - q_v^T) J_1(k) e^{-Kb} (\Lambda p_s + q_s) + \frac{1}{4} (p_v^T \Lambda + q_v^T) J_1(K) e^{-kb} (\Lambda p_s - q_s) \\
& - \frac{1}{4} p_v^T [(K+k) Q^+ - (K-k) Q^-] q_s + \frac{1}{4} q_v^T [(K+k) Q^+ + (K-k) Q^-] p_s \\
& + \frac{1}{2} (p_v^T \Lambda - q_v^T) [J_1(k) e^{-Kb} - Q^+] (\Lambda p_s + q_s) \\
& + \frac{1}{2} (p_v^T \Lambda + q_v^T) (K+\Lambda) D^-(\Lambda p_s + q_s) \\
& - \frac{1}{2} q_v^T (K+\Lambda) (Q^+ + Q^-) p_s + \frac{1}{2} p_v^T (K+\Lambda) (Q^+ - Q^-) q_s \\
& = \frac{1}{4} (p_v^T \Lambda + q_v^T) (K+k+2\Lambda) D^-(\Lambda p_s + q_s) \\
& + \frac{1}{4} (p_v^T \Lambda - q_v^T) J_1(k) e^{-Kb} (\Lambda p_s + q_s) + \frac{1}{4} (p_v^T \Lambda + q_v^T) J_1(K) e^{-kb} (\Lambda p_s - q_s) \\
& + \frac{1}{4} Q^+ [(K-k) (p_v^T q_s - q_v^T p_s) - 2p_v^T \Lambda^2 p_s + 2q_v^T q_s] \\
& - \frac{1}{4} Q^- [p_v^T (K+k+2\Lambda) q_s + q_v^T (K+k+2\Lambda) p_s] .
\end{aligned}$$

Similarly, for Eq. (A.34) one finds

$$\begin{aligned}
& \frac{1}{4} [h_2^{-T} g_1^- - h_1^{-T} g_2^-] + t_3^- + t_4^- = \frac{1}{4} (p_v^T \Lambda + q_v^T) (K-k) D^-(\Lambda p_s + q_s) \\
& - \frac{1}{4} (p_v^T \Lambda - q_v^T) J_1(k) e^{-Kb} (\Lambda p_s + q_s) + \frac{1}{4} (p_v^T \Lambda + q_v^T) J_1(K) e^{-kb} (\Lambda p_s - q_s) \\
& - \frac{1}{4} p_v^T [(K+k) Q^+ + (K-k) Q^-] q_s + \frac{1}{4} q_v^T [(K+k) Q^+ - (K-k) Q^-] p_s \\
& + \frac{1}{2} (p_v^T \Lambda - q_v^T) [J_1(k) e^{-Kb} - Q^+] (\Lambda p_s + q_s) \\
& - \frac{1}{2} (p_v^T \Lambda + q_v^T) (K+\Lambda) D^-(\Lambda p_s + q_s) \\
& - \frac{1}{2} q_v^T (K+\Lambda) (Q^+ - Q^-) p_s + \frac{1}{2} p_v^T (K+\Lambda) (Q^+ + Q^-) q_s \\
& = -\frac{1}{4} (p_v^T \Lambda + q_v^T) (K+k+2\Lambda) D^-(\Lambda p_s + q_s) \\
& + \frac{1}{4} (p_v^T \Lambda - q_v^T) J_1(k) e^{-Kb} (\Lambda p_s + q_s) + \frac{1}{4} (p_v^T \Lambda + q_v^T) J_1(K) e^{-kb} (\Lambda p_s - q_s) \\
& + \frac{1}{4} Q^+ [(K-k) (p_v^T q_s - q_v^T p_s) - 2p_v^T \Lambda^2 p_s + 2q_v^T q_s] \\
& + \frac{1}{4} Q^- [p_v^T (K+k+2\Lambda) q_s + q_v^T (K+k+2\Lambda) p_s] .
\end{aligned}$$

In both cases the result is completely symmetrical with respect to the directions of specular incidence and observation, by which the reciprocity relations for ρ_{so}^d and τ_{so}^d have been proved. In the expressions for the sum $\rho_{so}^d + \tau_{so}^d$ and the difference $\rho_{so}^d - \tau_{so}^d$ several terms differ only in sign, which makes it possible to obtain less complex expressions for ρ_{so}^d and τ_{so}^d separately. Writing



$$a^+ = \frac{1}{4} [h_2^{+T} Y^{-1} + h_1^{+T} Y^T] (R+T) [Y g_1^+ + Y^{-T} g_2^+] \\ + \frac{1}{4} [h_2^{+T} Y^{-1} Y^{-T} g_2^+ - h_1^{+T} Y^T Y g_1^+] \quad , \text{ and}$$

$$a^- = \frac{1}{4} [h_2^{-T} Y^{-1} + h_1^{-T} Y^T] (R-T) [Y g_1^- + Y^{-T} g_2^-] \\ + \frac{1}{4} [h_2^{-T} Y^{-1} Y^{-T} g_2^- - h_1^{-T} Y^T Y g_1^-] \quad .$$

the final result can be expressed as:

$$\rho_{so}^d = \frac{1}{2} (a^+ + a^-) + \frac{1}{4} (\Lambda p_v - q_v)^T J_1(k) e^{-Kb} (\Lambda p_s + q_s) \\ + \frac{1}{4} (\Lambda p_v + q_v)^T J_1(K) e^{-kb} (\Lambda p_s - q_s) \\ + \frac{1}{4} Q^+ [(K-k) (p_v^T q_s - q_v^T p_s) - 2p_v^T \Lambda^2 p_s + 2q_v^T q_s] \quad , \text{ and} \quad (A.35)$$

$$r_{so}^d = \frac{1}{2} (a^+ - a^-) + \frac{1}{4} (\Lambda p_v + q_v)^T (K+k+2\Lambda) D^- (\Lambda p_s + q_s) \\ + \frac{1}{4} Q^- [p_v^T (K+k+2\Lambda) q_s + q_v^T (K+k+2\Lambda) p_s] \quad . \quad (A.36)$$

In the derivation of the directional reflectances and transmittances, which was based on the integration of the functions $v^T E^- + v'^T E^+$ and $v'^T E^- + v^T E^+$, the azimuth transformation by means of the discrete cosine transform has not been mentioned, because it is not relevant for the formal mathematical results. But in practice it is much more efficient to describe the diffuse fluxes by their azimuthal transforms, E_m^- and E_m^+ , where the mode m varies from zero to six (in the example), and the transformed vectors E_m^- and E_m^+ have only three elements instead of 36. In this case the inner product $v^T E$ is written as a summation over all azimuthal modes by

$$v^T E = \sum_{m=0}^6 v_m^T E_m \quad ,$$

where v_m is a vector of scattering coefficients of three elements, which have been transformed by the inverse DCT. For the contribution to $v^T E$ from a single zenith zone one can write



$$c = v_1 E_1 + \dots + v_{12} E_{12} \quad ,$$

where E_1 to E_{12} are the fluxes in the twelve azimuthal sectors, and v_1 to v_{12} are the associate scattering coefficients. Because of the symmetry with respect to the principal plane, this can be written as

$$\begin{aligned} c &= v_1 E_1 + (v_2 + v_{12}) E_2 + \dots + (v_6 + v_8) E_6 + v_7 E_7 \\ &= E_1 v_1 + E_2 (v_2 + v_{12}) + \dots + E_6 (v_6 + v_8) + E_7 v_7 \quad , \text{ or} \\ c &= E^T v \quad . \end{aligned}$$

Since the azimuthally transformed flux vector $E' = WE$, and $W^2 = 12I$, $WE' = 12E$, or $12E^T = E'^T W^T$. This gives

$$c = E'^T \left(\frac{1}{12} W^T v \right) = E'^T v' \quad .$$

From this it is concluded that incorporation of the inverse DCT for the vector of azimuthal scattering coefficients v consists of applying the matrix $(1/12)W^T$ to a vector which consists of the elements v_1 , $(v_2 + v_{12})$, ..., $(v_6 + v_8)$ and v_7 . This transformation must be executed for each zenith zone and the results are placed in vectors v_m .

Equations (A.14) to (A.17), which give $T+R$, $T-R$, $\tau_{sd} + \rho_{sd}$ and $\tau_{sd} - \rho_{sd}$, (A.29) and (A.30), which give $\rho_{do}^T + \tau_{do}^T$ and $\rho_{do}^T - \tau_{do}^T$, and (A.35) and (A.36), which give ρ_{so}^d and τ_{so}^d , are all applicable for each mode m , and the inverse DCT must be applied in order to express the results in untransformed incident diffuse fluxes. For ρ_{so}^d and τ_{so}^d this boils down to a summation over all modes, so that one can write

$$\begin{aligned} \rho_{so} &= \rho_{so}^s + \sum_{m=0}^6 \rho_{so}^d(m) \quad , \text{ and} \\ \tau_{so} &= \tau_{so}^s + \sum_{m=0}^6 \tau_{so}^d(m) \quad , \end{aligned}$$

where ρ_{so}^s and τ_{so}^s are the single scattering contributions given by $\rho_{so}^s = w Q^+$ and $\tau_{so}^s = w' Q^-$.



Appendix B Approximations for small and infinite optical thickness

I. Small optical thickness

The functions depending on the optical thickness b are approximated as follows:

$$\mathbf{C}_\epsilon = \mathbf{I} + \mathbf{I} - \Lambda b = 2\mathbf{I} - \Lambda b \quad ;$$

$$\mathbf{S}_\epsilon = \Lambda^{-1} [\mathbf{I} - (\mathbf{I} - \Lambda b + \frac{1}{2} \Lambda^2 b^2)] = b [\mathbf{I} - \frac{1}{2} \Lambda b] \quad .$$

From this it follows that $(\mathbf{C}^{-1} \mathbf{S})_\epsilon = \frac{1}{2} b \mathbf{I} \quad .$

$$\begin{aligned} J_1(k)_\epsilon &= (k - \Lambda)^{-1} [(\mathbf{I} - \Lambda b + \frac{1}{2} \Lambda^2 b^2) - (\mathbf{I} - kb + \frac{1}{2} k^2 b^2)] \\ &= b [\mathbf{I} - \frac{1}{2} (k + \Lambda) b] \quad ; \end{aligned}$$

$$J_1(K)_\epsilon = b [\mathbf{I} - \frac{1}{2} (K + \Lambda) b] \quad ;$$

$$\mathbf{D}_\epsilon^- = (K - k)^{-1} [J_1(k)_\epsilon - J_1(K)_\epsilon] = \frac{1}{2} b^2 \mathbf{I} \quad ;$$

$$[J_1(k) e^{-Kb}]_\epsilon = b [\mathbf{I} - \frac{1}{2} (2K + k + \Lambda) b] \quad ;$$

$$[J_1(K) e^{-kb}]_\epsilon = b [\mathbf{I} - \frac{1}{2} (2k + K + \Lambda) b] \quad ;$$

$$\mathcal{Q}_\epsilon^+ = b [1 - \frac{1}{2} (K + k) b] \quad ;$$

$$\mathcal{Q}_\epsilon^- = b [1 - \frac{1}{2} (K + k) b] \quad ;$$

From $\mathbf{C}_\epsilon$ and $\mathbf{S}_\epsilon$ one finds $(\mathbf{R} + \mathbf{T})_\epsilon$ and $(\mathbf{R} - \mathbf{T})_\epsilon$ as follows:

$$\begin{aligned} \mathbf{R} + \mathbf{T} &= [\mathbf{C} \mathbf{Y}^T + \mathbf{S} \Lambda^2 \mathbf{Y}^{-1}]^{-1} [\mathbf{C} \mathbf{Y}^T - \mathbf{S} \Lambda^2 \mathbf{Y}^{-1}] \\ &= [\mathbf{I} + \mathbf{Y}^{-T} \mathbf{C}^{-1} \mathbf{S} \Lambda^2 \mathbf{Y}^{-1}]^{-1} [\mathbf{I} - \mathbf{Y}^{-T} \mathbf{C}^{-1} \mathbf{S} \Lambda^2 \mathbf{Y}^{-1}] \end{aligned}$$

$$\begin{aligned} \text{so } (\mathbf{R} + \mathbf{T})_\epsilon &= [\mathbf{I} + \mathbf{Y}^{-T} \frac{1}{2} b \mathbf{I} \Lambda^2 \mathbf{Y}^{-1}]^{-1} [\mathbf{I} - \mathbf{Y}^{-T} \frac{1}{2} b \mathbf{I} \Lambda^2 \mathbf{Y}^{-1}] \\ &= [\mathbf{I} + \frac{1}{2} b \mathbf{Y}^{-T} \Lambda^2 \mathbf{Y}^{-1}]^{-1} [\mathbf{I} - \frac{1}{2} b \mathbf{Y}^{-T} \Lambda^2 \mathbf{Y}^{-1}] \\ &\approx \mathbf{I} - b \mathbf{Y}^{-T} \Lambda^2 \mathbf{Y}^{-1} = \mathbf{I} - b(\mathbf{A} - \mathbf{B}) \quad . \end{aligned}$$



$$\begin{aligned} R - T &= [SY^T + CY^{-1}]^{-1} [SY^T - CY^{-1}] \\ &= [YC^{-1}SY^T + I]^{-1} [YC^{-1}SY^T - I] \end{aligned}$$

$$\begin{aligned} \text{so } (R - T)_\epsilon &= [Y^{1/2}bIY^T + I]^{-1} [Y^{1/2}bIY^T - I] \\ &= [^{1/2}bYY^T + I]^{-1} [^{1/2}bYY^T - I] \\ &\approx -I + bYY^T = -I + b(A + B) \quad . \end{aligned}$$

By means of the g -vectors, first order approximations appear to be sufficient in order to find $(\rho_{sd} + \tau_{sd})_\epsilon$ and $(\rho_{sd} - \tau_{sd})_\epsilon$. In this case the g -vectors are approximated by

$$\begin{aligned} (g_1^+)_\epsilon &= (2 - kb)p_s - b(\Lambda p_s + q_s) = 2p_s - b[(k + \Lambda)p_s + q_s] \quad ; \\ (g_2^+)_\epsilon &= kbq_s + \Lambda b(\Lambda p_s + q_s) = b[(k + \Lambda)q_s + \Lambda^2 p_s] \quad ; \\ (g_1^-)_\epsilon &= kb p_s + b(\Lambda p_s + q_s) = b[(k + \Lambda)p_s + q_s] \quad ; \\ (g_2^-)_\epsilon &= (2 - kb)q_s - \Lambda b(\Lambda p_s + q_s) = 2q_s - b[(k + \Lambda)q_s + \Lambda^2 p_s] \quad . \end{aligned}$$

Now,

$$\begin{aligned} \rho_{sd} + \tau_{sd} &= [CY^T + S\Lambda^2 Y^{-1}]^{-1} [Cg_2^+ - S\Lambda^2 g_1^+] \\ &= [I + Y^{-T}C^{-1}S\Lambda^2 Y^{-1}]^{-1} [Y^{-T}g_2^+ - Y^{-T}C^{-1}S\Lambda^2 g_1^+] \end{aligned}$$

$$\begin{aligned} \text{so } (\rho_{sd} + \tau_{sd})_\epsilon &\approx [I + ^{1/2}b(A - B)]^{-1} [Y^{-T}g_2^+ - ^{1/2}bY^{-T}\Lambda^2 g_1^+]_\epsilon \\ &\approx bY^{-T}[(k + \Lambda)q_s + \Lambda^2 p_s] - bY^{-T}\Lambda^2 [p_s - ^{1/2}b(k + \Lambda)p_s - ^{1/2}bq_s] \\ &\approx bY^{-T}(k + \Lambda)q_s = b(s' + s) \quad . \end{aligned}$$

Similarly,



$$\begin{aligned}\rho_{sd} - \tau_{sd} &= [\mathbf{S}\mathbf{Y}^T + \mathbf{C}\mathbf{Y}^{-1}]^{-1} [\mathbf{S}\mathbf{g}_2^- - \mathbf{C}\mathbf{g}_1^-] \\ &= [\mathbf{Y}\mathbf{C}^{-1}\mathbf{S}\mathbf{Y}^T + \mathbf{I}]^{-1} [\mathbf{Y}\mathbf{C}^{-1}\mathbf{S}\mathbf{g}_2^- - \mathbf{Y}\mathbf{g}_1^-]\end{aligned}$$

$$\begin{aligned}\text{so } (\rho_{sd} - \tau_{sd})_\epsilon &\approx [\mathbf{I} + \frac{1}{2}b(\mathbf{A} + \mathbf{B})]^{-1} [\frac{1}{2}b\mathbf{Y}\mathbf{g}_2^- - \mathbf{Y}\mathbf{g}_1^-]_\epsilon \\ &\approx b\mathbf{Y}[\mathbf{q}_s - \frac{1}{2}b(k + \Lambda)\mathbf{q}_s - \frac{1}{2}b\Lambda^2\mathbf{p}_s] - b\mathbf{Y}[(k + \Lambda)\mathbf{p}_s + \mathbf{q}_s] \\ &\approx -b\mathbf{Y}(k + \Lambda)\mathbf{p}_s = -b(s' - s) = b(s - s')\end{aligned}$$

Approximations up to the second order in b are necessary in order to analyze the behaviour of ρ_{so}^d and τ_{so}^d for small b . For the $\mathbf{g}$ -vectors and the $\mathbf{h}$ -vectors this gives

$$\begin{aligned}(\mathbf{g}_1^+)_\epsilon &= 2\mathbf{p}_s - b[(k + \Lambda)\mathbf{p}_s + \mathbf{q}_s] + \frac{1}{2}b^2[(k^2 + k\Lambda + \Lambda^2)\mathbf{p}_s + (k + \Lambda)\mathbf{q}_s] ; \\ (\mathbf{g}_2^+)_\epsilon &= b[(k + \Lambda)\mathbf{q}_s + \Lambda^2\mathbf{p}_s] - \frac{1}{2}b^2[(k^2 + k\Lambda + \Lambda^2)\mathbf{q}_s + (k + \Lambda)\Lambda^2\mathbf{p}_s] ; \\ (\mathbf{g}_1^-)_\epsilon &= b[(k + \Lambda)\mathbf{p}_s + \mathbf{q}_s] - \frac{1}{2}b^2[(k^2 + k\Lambda + \Lambda^2)\mathbf{p}_s + (k + \Lambda)\mathbf{q}_s] ; \\ (\mathbf{g}_2^-)_\epsilon &= 2\mathbf{q}_s - b[(k + \Lambda)\mathbf{q}_s + \Lambda^2\mathbf{p}_s] + \frac{1}{2}b^2[(k^2 + k\Lambda + \Lambda^2)\mathbf{q}_s + (k + \Lambda)\Lambda^2\mathbf{p}_s] ; \\ (\mathbf{h}_1^+)_\epsilon &= 2\mathbf{p}_v - b[(K + \Lambda)\mathbf{p}_v + \mathbf{q}_v] + \frac{1}{2}b^2[(K^2 + K\Lambda + \Lambda^2)\mathbf{p}_v + (K + \Lambda)\mathbf{q}_v] ; \\ (\mathbf{h}_2^+)_\epsilon &= b[(K + \Lambda)\mathbf{q}_v + \Lambda^2\mathbf{p}_v] - \frac{1}{2}b^2[(K^2 + K\Lambda + \Lambda^2)\mathbf{q}_v + (K + \Lambda)\Lambda^2\mathbf{p}_v] ; \\ (\mathbf{h}_1^-)_\epsilon &= b[(K + \Lambda)\mathbf{p}_v + \mathbf{q}_v] - \frac{1}{2}b^2[(K^2 + K\Lambda + \Lambda^2)\mathbf{p}_v + (K + \Lambda)\mathbf{q}_v] ; \\ (\mathbf{h}_2^-)_\epsilon &= 2\mathbf{q}_v - b[(K + \Lambda)\mathbf{q}_v + \Lambda^2\mathbf{p}_v] + \frac{1}{2}b^2[(K^2 + K\Lambda + \Lambda^2)\mathbf{q}_v + (K + \Lambda)\Lambda^2\mathbf{p}_v] .\end{aligned}$$

From the equations for ρ_{so}^d and τ_{so}^d one can write

$$\begin{aligned}4(\rho_{so}^d)_\epsilon &= 2(a^+ + a^-)_\epsilon + (\mathbf{p}_v^T \Lambda - \mathbf{q}_v^T) b [\mathbf{I} - \frac{1}{2}b(2\mathbf{K} + \mathbf{k} + \Lambda)] (\Lambda\mathbf{p}_s + \mathbf{q}_s) \\ &\quad + (\mathbf{p}_v^T \Lambda + \mathbf{q}_v^T) b [\mathbf{I} - \frac{1}{2}b(2\mathbf{k} + \mathbf{K} + \Lambda)] (\Lambda\mathbf{p}_s - \mathbf{q}_s) \\ &\quad + b [1 - \frac{1}{2}b(\mathbf{K} + \mathbf{k})] [(\mathbf{K} - \mathbf{k})(\mathbf{p}_v^T \mathbf{q}_s - \mathbf{q}_v^T \mathbf{p}_s) - 2\mathbf{p}_v^T \Lambda^2 \mathbf{p}_s + 2\mathbf{q}_v^T \mathbf{q}_s] , \text{ and}\end{aligned}$$

$$4(\tau_{so}^d)_\epsilon = 2(a^+ - a^-)_\epsilon + (p_v^T \Lambda + q_v^T)(K+k+2\Lambda) \frac{1}{2} b^2 I(\Lambda p_s + q_s) \\ - b[1 - \frac{1}{2} b(K+k)] [p_v^T(K+k+2\Lambda)q_s + q_v^T(K+k+2\Lambda)p_s] , \text{ giving}$$

$$4(\rho_{so}^d + \tau_{so}^d)_\epsilon = 4a^+_\epsilon + p_v^T [-2b(k+\Lambda) + b^2(K+k+\Lambda)(k+\Lambda)] q_s \\ + q_v^T [-2b(K+\Lambda) + b^2(K+k+\Lambda)(K+\Lambda)] p_s \\ + q_v^T b^2(K+k+2\Lambda) q_s , \text{ and}$$

$$4(\rho_{so}^d - \tau_{so}^d)_\epsilon = 4a^-_\epsilon + p_v^T [2b(K+\Lambda) - b^2(K+k+\Lambda)(K+\Lambda)] q_s \\ + q_v^T [2b(k+\Lambda) - b^2(K+k+\Lambda)(k+\Lambda)] p_s \\ - p_v^T b^2(K+k+2\Lambda)\Lambda^2 p_s .$$

Using $(R+T)_\epsilon$ in the expression for a^+_ϵ gives

$$4a^+_\epsilon = [h_2^{+T} Y^{-1} + h_1^{+T} Y^T]_\epsilon [I - bY^{-T}\Lambda^2 Y^{-1}] [Y^{-T}g_2^+ + Yg_1^+]_\epsilon \\ + [h_2^{+T} Y^{-1} Y^{-T}g_2^+ - h_1^{+T} Y^T Yg_1^+]_\epsilon \\ = [2h_2^{+T} Y^{-1} Y^{-T}g_2^+ + h_2^{+T} g_1^+ + h_1^{+T} g_2^+]_\epsilon \\ - b[h_2^{+T} Y^{-1} Y^{-T}\Lambda^2 Y^{-1} Y^{-T}g_2^+ + h_2^{+T} Y^{-1} Y^{-T}\Lambda^2 g_1^+ + h_1^{+T} \Lambda^2 Y^{-1} Y^{-T}g_2^+ + h_1^{+T} \Lambda^2 g_1^+]_\epsilon .$$

From this result, the term $[-bh_2^{+T} Y^{-1} Y^{-T}\Lambda^2 Y^{-1} Y^{-T}g_2^+]_\epsilon$ is neglected, since its lowest order contribution is proportional to b^3 . Substitution of the second-order approximations of the g -vectors and h -vectors and retaining only the terms up to the second order in b gives as a result

$$4a^+_\epsilon = 2b^2 [q_v^T(K+\Lambda) + p_v^T \Lambda^2] Y^{-1} Y^{-T} [(k+\Lambda)q_s + \Lambda^2 p_s] \\ + 2b [q_v^T(K+\Lambda) + p_v^T \Lambda^2] p_s + 2b p_v^T [(k+\Lambda)q_s + \Lambda^2 p_s] \\ - b^2 \{ [q_v^T(K^2 + K\Lambda + \Lambda^2) + p_v^T \Lambda^2(K+\Lambda)] p_s + [q_v^T(K+\Lambda) + p_v^T \Lambda^2] [(k+\Lambda)p_s + q_s] \} \\ - b^2 \{ p_v^T [(k^2 + k\Lambda + \Lambda^2)q_s + (k+\Lambda)\Lambda^2 p_s] + [p_v^T(K+\Lambda) + q_v^T] [(k+\Lambda)q_s + \Lambda^2 p_s] \} \\ - 2b^2 [q_v^T(K+\Lambda) + p_v^T \Lambda^2] Y^{-1} Y^{-T}\Lambda^2 p_s - 2b^2 p_v^T \Lambda^2 Y^{-1} Y^{-T} [(k+\Lambda)q_s + \Lambda^2 p_s] \\ - 4b p_v^T \Lambda^2 p_s + 2b^2 \{ p_v^T \Lambda^2 [(k+\Lambda)p_s + q_s] + [p_v^T(K+\Lambda) + q_v^T] \Lambda^2 p_s \} ,$$



which can be reduced to

$$4a^+_{\epsilon} = 2b [q_v^T(K+\Lambda)p_s + p_v^T(k+\Lambda)q_s] \\ + b^2 \{ 2q_v^T(K+\Lambda)Y^{-1}Y^{-T}(k+\Lambda)q_s - p_v^T(K+k+\Lambda)(k+\Lambda)q_s \\ - q_v^T(K+k+\Lambda)(K+\Lambda)p_s - q_v^T(K+k+2\Lambda)q_s \} \quad .$$

Using $(R-T)_{\epsilon}$ in a similar elaboration for a^-_{ϵ} gives

$$4a^-_{\epsilon} = [h_2^{-T}Y^{-1} + h_1^{-T}Y^T]_{\epsilon} [-I + bYY^T] [Y^{-T}g_2^- + Yg_1^-]_{\epsilon} \\ + [h_2^{-T}Y^{-1}Y^{-T}g_2^- - h_1^{-T}Y^TYg_1^-]_{\epsilon} \\ = [-2h_1^{-T}Y^TYg_1^- - h_2^{-T}g_1^- - h_1^{-T}g_2^-]_{\epsilon} \\ + b [h_2^{-T}g_2^- + h_2^{-T}Y^TYg_1^- + h_1^{-T}Y^TYg_2^- + h_1^{-T}Y^TYY^TYg_1^-]_{\epsilon} \quad .$$

From this, the term $[b h_1^{-T}Y^TYg_1^-]_{\epsilon}$ is neglected, since it is proportional to b^3 . Substitution of the second order approximations of the g -vectors and h -vectors and retaining only the terms up to the second order in this case gives

$$4a^-_{\epsilon} = -2b^2 [p_v^T(K+\Lambda) + q_v^T] Y^T Y [(k+\Lambda)p_s + q_s] \\ - 2b [p_v^T(K+\Lambda) + q_v^T] q_s - 2b q_v^T [(k+\Lambda)p_s + q_s] \\ + b^2 \{ [p_v^T(K^2 + K\Lambda + \Lambda^2) + q_v^T(K+\Lambda)] q_s + [p_v^T(K+\Lambda) + q_v^T] [(k+\Lambda)q_s + \Lambda^2 p_s] \} \\ + b^2 \{ q_v^T [(k^2 + k\Lambda + \Lambda^2)p_s + (k+\Lambda)q_s] + [q_v^T(K+\Lambda) + p_v^T\Lambda^2] [(k+\Lambda)p_s + q_s] \} \\ + 2b^2 [p_v^T(K+\Lambda) + q_v^T] Y^T Y q_s + 2b^2 q_v^T Y^T Y [(k+\Lambda)p_s + q_s] \\ + 4b q_v^T q_s - 2b^2 \{ q_v^T [(k+\Lambda)q_s + \Lambda^2 p_s] + [q_v^T(K+\Lambda) + p_v^T\Lambda^2] q_s \} \quad ,$$

which reduces to

$$4a^-_{\epsilon} = -2b [q_v^T(k+\Lambda)p_s + p_v^T(K+\Lambda)q_s] \\ + b^2 \{ -2p_v^T(K+\Lambda)Y^T Y(k+\Lambda)p_s + q_v^T(K+k+\Lambda)(k+\Lambda)p_s \\ + p_v^T(K+k+\Lambda)(K+\Lambda)q_s + p_v^T\Lambda^2(K+k+2\Lambda)p_s \} \quad .$$

Finally, substitution of a^+_{ϵ} and a^-_{ϵ} in the expressions for $(\rho_{so}^d + \tau_{so}^d)_{\epsilon}$ and $(\rho_{so}^d - \tau_{so}^d)_{\epsilon}$ gives

$$4(\rho_{so}^d + \tau_{so}^d)_\epsilon = 2b^2 \mathbf{q}_v^T (K + \Lambda) Y^{-1} Y^{-T} (k + \Lambda) \mathbf{q}_s, \text{ and}$$

$$4(\rho_{so}^d - \tau_{so}^d)_\epsilon = -2b^2 \mathbf{p}_v^T (K + \Lambda) Y^T Y (k + \Lambda) \mathbf{p}_s.$$

Substitution of $\mathbf{p}_v^T$, $\mathbf{q}_v^T$, $\mathbf{p}_s$ and $\mathbf{q}_s$ in these expression gives

$$(\rho_{so}^d + \tau_{so}^d)_\epsilon = \frac{1}{2} b^2 (\mathbf{v}'^T + \mathbf{v}^T) (\mathbf{s}' + \mathbf{s}), \text{ and}$$

$$(\rho_{so}^d - \tau_{so}^d)_\epsilon = -\frac{1}{2} b^2 (\mathbf{v}'^T - \mathbf{v}^T) (\mathbf{s}' - \mathbf{s}).$$

From the sum and difference of the above equations one obtains

$$(\rho_{so}^d)_\epsilon = \frac{1}{2} b^2 (\mathbf{v}'^T \mathbf{s} + \mathbf{v}^T \mathbf{s}'), \text{ and}$$

$$(\tau_{so}^d)_\epsilon = \frac{1}{2} b^2 (\mathbf{v}'^T \mathbf{s}' + \mathbf{v}^T \mathbf{s}).$$

II Infinite optical thickness

In this case the functions depending on the optical thickness b are approximated by:

$$\begin{aligned} \mathbf{C}_\infty &= \mathbf{I} & ; & \quad \mathbf{S}_\infty = \Lambda^{-1} & ; \\ \mathbf{J}_1(k)_\infty &= \mathbf{O} & ; & \quad \mathbf{J}_1(K)_\infty = \mathbf{O} & ; & \quad \mathbf{D}_\infty^- = \mathbf{O} & ; \\ \mathbf{Q}_\infty^- &= 0 & ; & \quad \mathbf{Q}_\infty^+ = (K + k)^{-1} & . \end{aligned}$$

As the exponential functions e^{-kb} and e^{-Kb} also become zero for infinite optical thickness, the $\mathbf{g}$ -vectors and $\mathbf{h}$ -vectors are approximated by

$$\begin{aligned} (\mathbf{g}_1^+)_\infty &= \mathbf{p}_s & ; & \quad (\mathbf{g}_2^+)_\infty = \mathbf{q}_s & ; & \quad (\mathbf{g}_1^-)_\infty = \mathbf{p}_s & ; & \quad (\mathbf{g}_2^-)_\infty = \mathbf{q}_s & ; \\ (\mathbf{h}_1^+)_\infty &= \mathbf{p}_v & ; & \quad (\mathbf{h}_2^+)_\infty = \mathbf{q}_v & ; & \quad (\mathbf{h}_1^-)_\infty = \mathbf{p}_v & ; & \quad (\mathbf{h}_2^-)_\infty = \mathbf{q}_v & . \end{aligned}$$

With the above approximations one obtains



$$\begin{aligned}
 (R+T)_\infty &= [Y^T + \Lambda Y^{-1}]^{-1} [Y^T - \Lambda Y^{-1}] \quad ; \\
 (R-T)_\infty &= [\Lambda^{-1} Y^T + Y^{-1}]^{-1} [\Lambda^{-1} Y^T - Y^{-1}] = [Y^T + \Lambda Y^{-1}]^{-1} [Y^T - \Lambda Y^{-1}] \\
 &= (R+T)_\infty \quad .
 \end{aligned}$$

From this it follows that $T_\infty = O$ and $R_\infty = [Y^T + \Lambda Y^{-1}]^{-1} [Y^T - \Lambda Y^{-1}]$.

For $(\rho_{sd} + \tau_{sd})_\infty$ and $(\rho_{sd} - \tau_{sd})_\infty$ one finds

$$\begin{aligned}
 (\rho_{sd} + \tau_{sd})_\infty &= [Y^T + \Lambda Y^{-1}]^{-1} [q_s - \Lambda p_s] \quad ; \\
 (\rho_{sd} - \tau_{sd})_\infty &= [\Lambda^{-1} Y^T + Y^{-1}]^{-1} [\Lambda^{-1} q_s - p_s] = [Y^T + \Lambda Y^{-1}]^{-1} [q_s - \Lambda p_s] \\
 &= (\rho_{sd} + \tau_{sd})_\infty \quad .
 \end{aligned}$$

From this it one obtains $(\tau_{sd})_\infty = 0$ and

$$(\rho_{sd})_\infty = [Y^T + \Lambda Y^{-1}]^{-1} [q_s - \Lambda p_s] \quad .$$

For the terms a^+ and a^- substitution of the g -vectors and h -vectors, and of $R+T$ and $R-T$ for $b = \infty$ gives

$$a^+_\infty = a^-_\infty = \frac{1}{4} (p_v^T Y^T + q_v^T Y^{-1}) R_\infty (Y p_s + Y^{-T} q_s) + \frac{1}{4} (q_v^T Y^{-1} Y^{-T} q_s - p_v^T Y^T Y p_s) \quad .$$

This gives $(\tau_{sd}^d)_\infty = 0$ and

$$(\rho_{so}^d)_\infty = a^+_\infty + \frac{1}{4} (K+k)^{-1} [(K-k) (p_v^T q_s - q_v^T p_s) - 2p_v^T \Lambda^2 p_s + 2q_v^T q_s] \quad .$$



Appendix C Extinction and scattering coefficients of the SAIL model

The extinction and scattering coefficients of the SAIL model are based on those for specular fluxes and a fixed leaf orientation.

In order to include also hemispherically diffuse (semi-isotropic) fluxes, first it is determined what portions of a downward semi-isotropic irradiance E^- are incident on both sides of a leaf inclined at an angle θ_l . The radiance of the upper hemisphere in this case is constant and equal to $L^- = E^-/\pi$. In Fig. C.1 it is illustrated that by means of a vector p , which is perpendicular to the vertical n and the leaf normal ℓ , spherical co-ordinates can be employed which facilitate the integration, namely a "zenith" angle α and an "azimuth" angle β . In this co-ordinate system the vector ℓ is given by $\alpha = \pi/2$ and $\beta = \pi/2 - \theta_l$. The irradiance contribution from an element of the

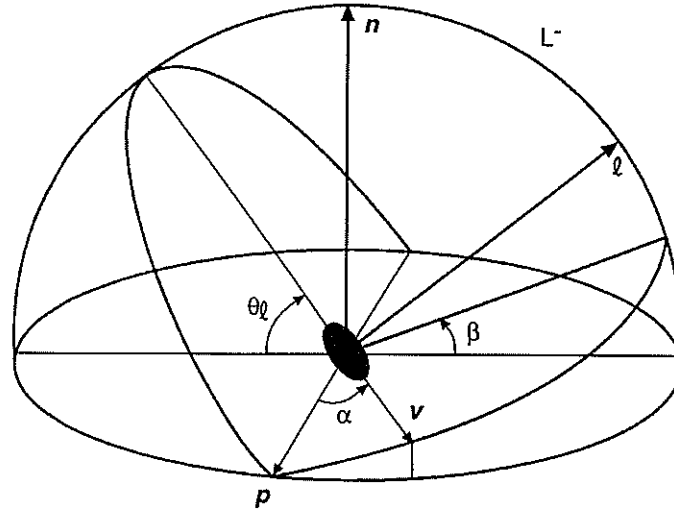


Fig. C.1 On the calculation of the diffuse irradiance on a leaf

upper hemisphere with solid angle $d\Omega = \sin \alpha d\alpha d\beta$ to the total irradiance on the leaf is given by

$$dE_i = L^- |\ell \cdot v| d\Omega ,$$

$$\text{where } \ell = (0 ; \sin \theta_l ; \cos \theta_l) ,$$

$$\text{and } v = (\cos \alpha ; \sin \alpha \cos \beta ; \sin \alpha \sin \beta) .$$

For the inner product this gives $\ell \cdot v = \sin \alpha [\sin \theta_l \cos \beta + \cos \theta_l \sin \beta] =$



$\sin \alpha \sin (\beta + \theta_i)$, so that

$$dE_i = L^- | \sin^2 \alpha d\alpha \sin (\beta + \theta_i) d\beta | .$$

The absolute value must be taken since the inner product $\ell \cdot \nu$ becomes negative for $\beta + \theta_i > \pi$. In that case the lower side of the leaf is illuminated. Therefore, the integration is carried out separately for the upper and the lower side of the leaf. For the upper side of the leaf one finds

$$\begin{aligned} E_1 &= L^- \int_0^\pi \sin^2 \alpha d\alpha \int_0^{\pi-\theta_i} \sin (\beta + \theta_i) d\beta = L^- \cdot \frac{\pi}{2} [-\cos \pi + \cos \theta_i] \\ &= \pi L^- \frac{1 + \cos \theta_i}{2} . \end{aligned}$$

For the lower side of the leaf the result is found as

$$\begin{aligned} E_2 &= L^- \int_0^\pi \sin^2 \alpha d\alpha \int_{\pi-\theta_i}^\pi -\sin (\beta + \theta_i) d\beta = L^- \cdot \frac{\pi}{2} [\cos \pi - \cos (\pi + \theta_i)] \\ &= \pi L^- \frac{1 - \cos \theta_i}{2} . \end{aligned}$$

From these results it is concluded that $E_1 + E_2 = \pi L^- = E^-$, so the irradiance intercepted by an inclined leaf equals the total incident irradiance, regardless of the leaf inclination angle. The fractions incident on both sides are called $f_1 = (1 + \cos \theta_i)/2$ and $f_2 = (1 - \cos \theta_i)/2$, where f_1 is the greater of the two. The same fractions are obtained for the interception of upward diffuse incident flux and for the contributions to scattered diffuse fluxes if the radiances of both sides of the leaf are given.

Together with the factors f_s and f_o , which are introduced in chapter 7, the extinction and scattering coefficients for fixed leaf orientation are easily expressed as follows.

Extinction coefficients:

$$\begin{aligned} k(\theta_i, \varphi_i) &= |f_s| \\ K(\theta_i, \varphi_i) &= |f_o| \\ \kappa(\theta_i, \varphi_i) &= f_1 + f_2 = 1 \end{aligned}$$



Scattering coefficients:

$$\begin{aligned}
 w(\theta_t, \varphi_t) &= f_s \rho f_o & \text{or} & \quad -f_s \tau f_o \\
 s(\theta_t, \varphi_t) &= f_s (\rho f_1 + \tau f_2) & \text{or} & \quad -f_s (\tau f_1 + \rho f_2) \\
 s'(\theta_t, \varphi_t) &= f_s (\tau f_1 + \rho f_2) & \text{or} & \quad -f_s (\rho f_1 + \tau f_2) \\
 v(\theta_t, \varphi_t) &= (f_1 \rho + f_2 \tau) f_o & \text{or} & \quad -(f_1 \tau + f_2 \rho) f_o \\
 v'(\theta_t, \varphi_t) &= (f_1 \tau + f_2 \rho) f_o & \text{or} & \quad -(f_1 \rho + f_2 \tau) f_o \\
 \sigma(\theta_t, \varphi_t) &= f_1 (\rho f_1 + \tau f_2) + f_2 (\tau f_1 + \rho f_2) \\
 \sigma'(\theta_t, \varphi_t) &= f_1 (\tau f_1 + \rho f_2) + f_2 (\rho f_1 + \tau f_2)
 \end{aligned}$$

The SAIL coefficients are now obtained by averaging the above coefficients over the leaf azimuth angle φ_t , under the assumption of a uniform leaf azimuth distribution.

For the coefficients κ , σ and σ' the results are equal to those listed above, as the factors f_1 and f_2 are independent of the leaf azimuth, giving

$\kappa(\theta_t) = 1$, and taking sum and difference of σ and σ' gives

$$\begin{aligned}
 \sigma(\theta_t) + \sigma'(\theta_t) &= (f_1 + f_2)(\rho + \tau)(f_1 + f_2) = \rho + \tau, & \text{and} \\
 \sigma(\theta_t) - \sigma'(\theta_t) &= (f_1 - f_2)(\rho - \tau)(f_1 - f_2) = (\rho - \tau) \cos^2 \theta_t, & \text{so that} \\
 \sigma(\theta_t) &= \frac{1}{2} [\rho + \tau + (\rho - \tau) \cos^2 \theta_t], & \text{and} \\
 \sigma'(\theta_t) &= \frac{1}{2} [\rho + \tau - (\rho - \tau) \cos^2 \theta_t].
 \end{aligned}$$

For the azimuthal averages of the factors f_s and f_o , these are integrated separately over the positive and negative parts. The factors f_s and f_o are given by

$$\begin{aligned}
 f_s &= [C_s + S_s \cos \varphi_t] / \cos \theta_s, & \text{and} \\
 f_o &= [C_o + S_o \cos(\varphi_t - \psi)] / |\cos \theta_o|, \\
 \text{with } C_s &= \cos \theta_t \cos \theta_s; \quad S_s = \sin \theta_t \sin \theta_s; \\
 C_o &= \cos \theta_t \cos \theta_o; \quad S_o = \sin \theta_t \sin \theta_o.
 \end{aligned}$$

The factor f_s becomes negative if $S_s > C_s$ and $2\pi - \beta_s > \varphi_t > \beta_s$, where $\beta_s = \arccos(-C_s/S_s)$. This gives

$$k(\theta_t) = f_{st} + f_{sb} = \frac{2}{2\pi} \int_0^{\beta_s} f_s d\varphi_t + \frac{2}{2\pi} \int_{\beta_s}^{\pi} -f_s d\varphi_t \quad .$$

The factors f_{st} and f_{sb} are the contributions to the azimuthal average from leaves illuminated at the top and at the bottom, respectively, and these are given by

$$f_{st} = \frac{1}{\pi \cos \theta_s} [\beta_s C_s + S_s \sin \beta_s] \quad , \text{ and}$$

$$f_{sb} = \frac{1}{\pi \cos \theta_s} [(\beta_s - \pi) C_s + S_s \sin \beta_s] \quad .$$

From the sum of both the extinction coefficient $k(\theta_t)$ is found as

$$k(\theta_t) = \frac{2}{\pi} [(\beta_s - \pi/2) C_s + S_s \sin \beta_s] / \cos \theta_s \quad .$$

The azimuth angle β_s is equal to half the range of φ_t over which the top side of the leaf is illuminated. In case $S_s \leq C_s$, the top side is illuminated for any φ_t , so in this case β_s equals π , and one obtains $k(\theta_t) = C_s / \cos \theta_s = \cos \theta_t$.

Similar considerations for the factor f_o lead to the half azimuth range $\beta_o = \arccos(-C_o/S_o)$ over which the upper side of the leaf is observed. Again one may set $\beta_o = \pi$ if $S_o \leq C_o$. For the determination of the extinction coefficient in the viewing direction, $K(\theta_t)$, it is assumed that viewing is from above, so that $\theta_o < \pi/2$ and $C_o > 0$. In that case f_o is negative if $S_o > C_o$ and $2\pi - \beta_o > \varphi_t - \psi > \beta_o$.

The factors f_{ot} and f_{ob} are the contributions to the azimuthal average from leaves observed at the top and at the bottom, respectively, and are given by

$$f_{ot} = \frac{1}{\pi \cos \theta_o} [\beta_o C_o + S_o \sin \beta_o] \quad , \text{ and}$$

$$f_{ob} = \frac{1}{\pi \cos \theta_o} [(\beta_o - \pi) C_o + S_o \sin \beta_o] \quad .$$

From the sum of both the extinction coefficient $K(\theta_\ell)$ is found as

$$K(\theta_\ell) = \frac{2}{\pi} [(\beta_o - \pi/2) C_o + S_o \sin \beta_o] / \cos \theta_o .$$

The scattering coefficients $s(\theta_\ell)$, $s'(\theta_\ell)$, $v(\theta_\ell)$ and $v'(\theta_\ell)$ are easily obtained by making use of the factors f_{st} , f_{sb} , f_{ot} and f_{ob} , which were applied in the derivation of the extinction coefficients $k(\theta_\ell)$ and $K(\theta_\ell)$. This gives

$$\begin{aligned} s(\theta_\ell) &= f_{st}(\rho f_1 + \tau f_2) + f_{sb}(\tau f_1 + \rho f_2) , \\ s'(\theta_\ell) &= f_{st}(\tau f_1 + \rho f_2) + f_{sb}(\rho f_1 + \tau f_2) , \\ v(\theta_\ell) &= (f_1 \rho + f_2 \tau) f_{ot} + (f_1 \tau + f_2 \rho) f_{ob} , \\ v'(\theta_\ell) &= (f_1 \tau + f_2 \rho) f_{ot} + (f_1 \rho + f_2 \tau) f_{ob} . \end{aligned}$$

Taking sums and differences of $s(\theta_\ell)$ and $s'(\theta_\ell)$ gives

$$\begin{aligned} s(\theta_\ell) + s'(\theta_\ell) &= (f_{st} + f_{sb})(\rho + \tau)(f_1 + f_2) = (\rho + \tau)k(\theta_\ell) , \text{ and} \\ s(\theta_\ell) - s'(\theta_\ell) &= (f_{st} - f_{sb})(\rho - \tau)(f_1 - f_2) = (\rho - \tau)\cos^2 \theta_\ell , \text{ or} \end{aligned}$$

$$\begin{aligned} s(\theta_\ell) &= 1/2 [(\rho + \tau)k(\theta_\ell) + (\rho - \tau)\cos^2 \theta_\ell] , \text{ and} \\ s'(\theta_\ell) &= 1/2 [(\rho + \tau)k(\theta_\ell) - (\rho - \tau)\cos^2 \theta_\ell] . \end{aligned}$$

Similarly, for $v(\theta_\ell)$ and $v'(\theta_\ell)$ one obtains

$$\begin{aligned} v(\theta_\ell) + v'(\theta_\ell) &= (f_1 + f_2)(\rho + \tau)(f_{ot} + f_{ob}) = (\rho + \tau)K(\theta_\ell) , \text{ and} \\ v(\theta_\ell) - v'(\theta_\ell) &= (f_1 - f_2)(\rho - \tau)(f_{ot} - f_{ob}) = (\rho - \tau)\cos^2 \theta_\ell , \text{ or} \end{aligned}$$

$$\begin{aligned} v(\theta_\ell) &= 1/2 [(\rho + \tau)K(\theta_\ell) + (\rho - \tau)\cos^2 \theta_\ell] , \text{ and} \\ v'(\theta_\ell) &= 1/2 [(\rho + \tau)K(\theta_\ell) - (\rho - \tau)\cos^2 \theta_\ell] . \end{aligned}$$

The bidirectional scattering coefficient $w(\theta_\ell)$ is determined from the integral

$$w(\theta_\ell) = \frac{1}{2\pi} \int_0^{2\pi} w(\theta_\ell, \varphi_\ell) d\varphi_\ell ,$$

where $w(\theta_\ell, \varphi_\ell)$ equals $f_s \rho f_o$ if the product $f_s f_o$ is positive and is equal to $-f_s \tau f_o$ if $f_s f_o$ is negative. In the first case illumination and observation are



at the same side of the leaf, whereas in the second case they are at different sides of the leaf.

For the analytical integration the product $f_s f_o$ is written out as

$$\begin{aligned} f_s f_o &= [C_s + S_s \cos \varphi_\ell] [C_o + S_o \cos(\varphi_\ell - \psi)] / |\cos \theta_s \cos \theta_o| \\ &= \frac{C_s C_o + \frac{1}{2} S_s S_o \cos \psi + S_s C_o \cos \varphi_\ell + C_s S_o \cos(\varphi_\ell - \psi) + \frac{1}{2} S_s S_o \cos(2\varphi_\ell - \psi)}{|\cos \theta_s \cos \theta_o|} . \end{aligned}$$

The indefinite integral of the numerator is called $J(\varphi_\ell)$ and given by

$$\begin{aligned} J(\varphi_\ell) &= [C_s C_o + \frac{1}{2} S_s S_o \cos \psi] \varphi_\ell + S_s C_o \sin \varphi_\ell + C_s S_o \sin(\varphi_\ell - \psi) \\ &\quad + \frac{1}{4} S_s S_o \sin(2\varphi_\ell - \psi) . \end{aligned}$$

Now, let J^+ denote the part of $J(\varphi_\ell)$ over the ranges of φ_ℓ for which the product $f_s f_o$ is positive, and J^- the integral over the ranges of φ_ℓ for which it is negative. In that case one can write

$$w(\theta_\ell) |\cos \theta_s \cos \theta_o| = \frac{1}{2\pi} [\rho J^+ - \tau J^-] . \quad (C.1)$$

Since $J^+ + J^- = J(2\pi) - J(0) = 2\pi [C_s C_o + \frac{1}{2} S_s S_o \cos \psi]$, it is sufficient to determine only J^- by writing

$$w(\theta_\ell) |\cos \theta_s \cos \theta_o| = \frac{1}{2\pi} [\rho (J^+ + J^-) - (\rho + \tau) J^-] .$$

The calculation of J^- is difficult, because several different cases must be distinguished, since in general there are two ranges of φ_ℓ for which the product $f_s f_o$ is negative and the lower and upper boundaries of integration vary depending on the relative azimuth difference ψ . However, a detailed analysis has resulted in the conclusion that by means of two auxiliary azimuthal difference angles Δ_1 and Δ_2 the number of different cases can be reduced to three. The angles Δ_1 and Δ_2 are defined by

$$\Delta_1 = |\beta_s - \beta_o| \text{ and } \Delta_2 = \pi - |\beta_s + \beta_o - \pi| .$$

The different cases are found by placing the angles ψ , Δ_1 and Δ_2 in ascending order. Since Δ_2 is always greater than or equal to Δ_1 , three cases



remain, namely $\psi \leq \Delta_1$, $\Delta_1 < \psi < \Delta_2$ and $\psi \geq \Delta_2$. By assigning the angles β_1 , β_2 and β_3 to the minimum, the median and the maximum of the angles ψ , Δ_1 and Δ_2 it is possible to express the integral J^- in a single formula, which reads

$$J^- = \beta_2 T_1(\theta_t) - T_2(\theta_t) \quad ,$$

where $T_1(\theta_t) = 2 C_s C_o + S_s S_o \cos \psi$, and

$$T_2(\theta_t) = \sin \beta_2 [2 D_s D_o + S_s S_o \cos \beta_1 \cos \beta_3] \quad ,$$

in which $D_s = -C_s / \cos \beta_s$ and $D_o = -C_o / \cos \beta_o$.

Note that $D_s = S_s$ if $\beta_s < \pi$ and $D_s = C_s$ if $\beta_s = \pi$. Also, $D_o = S_o$ if $0 < \beta_o < \pi$, $D_o = C_o$ if $\beta_o = \pi$ and $D_o = -C_o$ if $\beta_o = 0$.

Since $J^+ + J^-$ appears to be equal to $\pi T_1(\theta_t)$, J^+ is given by

$$J^+ = (\pi - \beta_2) T_1(\theta_t) + T_2(\theta_t) \quad .$$

Application of Eq. (C.1) then gives

$$w(\theta_t) = \frac{1}{2\pi} [\rho F_1(\theta_t) + \tau F_2(\theta_t)] \quad ,$$

$$\text{where } F_1(\theta_t) = \frac{(\pi - \beta_2) T_1(\theta_t) + T_2(\theta_t)}{|\cos \theta_s \cos \theta_o|} \quad ,$$

$$\text{and } F_2(\theta_t) = \frac{-\beta_2 T_1(\theta_t) + T_2(\theta_t)}{|\cos \theta_s \cos \theta_o|} \quad .$$

This concludes the derivation of the extinction and scattering coefficients of the SAIL model.