Development of a Statistical Method for Estimating the Downward Longwave Radiation at the Surface from Satellite Observations

by

Hai-Tien Lee

Dissertation submitted to the Faculty of the Graduate School of The University of Maryland in partial fulfillment of the requirement for the degree of Doctor of Philosophy

1993

Advisory Committee:

Professor Robert G. Ellingson, Chairman/Advisor
Dr. Arnold Gruber
Professor Benjamin Kedem
Professor Rachel Pinker
Professor Owen Thompson
ABSTRACT

Title of Dissertation: Development of a Statistical Method for Estimating the Downward Longwave Radiation at the Surface from Satellite Observations

Hai-Tien Lee, Doctor of Philosophy, 1993

Dissertation directed by: Robert G. Ellingson, Professor
Department of Meteorology

A new technique was developed statistically for estimating the DLR using the HIRS/2 clear-column radiances from NOAA polar orbiting satellites. Simulated radiation data were used in the model development and validation. Two categories of regression equations were studied. They are the linear models and the emissivity/transmissivity-approach models. It was found that the emissivity/transmissivity models are superior than the linear ones in the aspects of the RMS errors of the DLR estimates as well as in the residual spread patterns. For both clear and cloudy conditions, when instrument noise is included, the regression RMS errors of the emissivity and transmissivity models range from about 4 to 9 Wm\(^{-2}\).

Sensitivity studies show that DLR is the most sensitive to the error in the cloud amount on average. Overall, the combined DLR errors, excluding the contribution of the error in the surface pressure estimation, range from about 7 to 12 Wm\(^{-2}\) when there are ±10\% uncertainties in cloud amounts and ±100 mb uncertainties in cloud base heights. When the cloud amount
uncertainties rise to 30%, the range of the combined DLR error ranges from about 10 to 25 Wm\(^{-2}\). The standard deviations of DLR differences found in the validations are about 9 Wm\(^{-2}\) with mean differences at about ±1 Wm\(^{-2}\).

For cloudy skies, cloud base height and the effective cloud amount are needed. The DLR estimated with different cloud data sources, however, has not been examined. Since the sensitivity study shows that the DLR error due to the uncertainty of the cloud amount usually dominates the other error components, it is believed that an accurate estimate of the cloud amount is more important for cloudy-sky DLR estimation.

Despite these limitations noted above, our analysis indicates that the physical statistical method developed in this study should provide estimates of the DLR to within about 10 Wm\(^{-2}\) RMS directly from the HIRS observations if the cloud properties can be accurately estimated. Furthermore, it is expected that the errors will be considerably smaller for monthly average, except where persistent extreme conditions occur.
DEDICATION

To my parents
ACKNOWLEDGMENT

Professor Robert G. Ellingson, my dissertation advisor, deserves the most thanks from me. During the entire time of my graduate study, he has always provided me the best possible environment to pursue my research interests. I thank him for the countless hours he has spent with me discussing my research, directing me to relevant literature, and critiquing drafts of my dissertation. I am grateful to his constant financial support that made my graduate study possible.

Thanks also go to Dr. Arnold Gruber, Prof. Benjamin Kedem, Prof. Rachel Pinker and Prof. Owen Thompson for serving on my dissertation committee, and for their careful reading of the dissertation and their helpful comments. I would like to thank David Yanuk, Dr. Shi-Keng Yang and Dr. Sisong Zhou for providing valuable information and suggestions. I also thank Xiaojing Zhou for preparing part of the data and the initial validation of the technique. Special thanks to Jim Cooper, Austin Conaty and Brian Doty for their assistance and maintenance on the computer facilities that kept the research works going smoothly. I would also like to thank Dr. George Huffman who guided me into the field of meteorology.

Finally, I would like to thank my wife Yuan-Kuen for her love, constant support and encouragement. After all, I owe a great deal of gratitude to my parents. Thus, I want to dedicate this dissertation to my parents for their love, support and inspiration during my pursuit of higher education.

This research was sponsored in part by the National Oceanic and Atmospheric Administration through the Cooperative Institute for Climate Studies and by the National Aeronautics and Space Administration through Grant NAGW-1471,S4.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter I Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II Theoretical Basis</td>
<td>7</td>
</tr>
<tr>
<td>Basic Equations</td>
<td>7</td>
</tr>
<tr>
<td>Conceptual Framework</td>
<td>9</td>
</tr>
<tr>
<td>Chapter III Data for Regression Analyses and Validation</td>
<td>24</td>
</tr>
<tr>
<td>Radiative Transfer Model</td>
<td>25</td>
</tr>
<tr>
<td>Atmospheric Profiles and Radiation Data</td>
<td>27</td>
</tr>
<tr>
<td>Chapter IV Model Development and Regression Analyses</td>
<td>35</td>
</tr>
<tr>
<td>Clear Sky Condition</td>
<td>37</td>
</tr>
<tr>
<td>Linear Models</td>
<td>37</td>
</tr>
<tr>
<td>Emissivity-approach Models</td>
<td>43</td>
</tr>
<tr>
<td>Cloudy Sky Condition</td>
<td>51</td>
</tr>
<tr>
<td>Linear Models</td>
<td>51</td>
</tr>
<tr>
<td>Transmissivity-approach Models</td>
<td>54</td>
</tr>
<tr>
<td>Topographical Adjustment</td>
<td>59</td>
</tr>
<tr>
<td>Chapter V Error Analysis and Sensitivity Study</td>
<td>87</td>
</tr>
<tr>
<td>General Discussion</td>
<td>87</td>
</tr>
<tr>
<td>Error Analysis</td>
<td>88</td>
</tr>
<tr>
<td>Sensitivity Study</td>
<td>91</td>
</tr>
<tr>
<td>Chapter VI Validation and Analysis</td>
<td>112</td>
</tr>
<tr>
<td>Data Description</td>
<td>112</td>
</tr>
<tr>
<td>Validation Results</td>
<td>113</td>
</tr>
<tr>
<td>Chapter VII Summary, Conclusions and Suggestions for Future Research</td>
<td>139</td>
</tr>
<tr>
<td>References</td>
<td>145</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Number</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 2.1.</strong> Specifications and typical characteristics of HIRS/2 on NOAA-10 satellite.</td>
<td>23</td>
</tr>
<tr>
<td><strong>Table 4.1.</strong> Regression coefficients and their standard errors for variables in Model I-1.</td>
<td>38</td>
</tr>
<tr>
<td><strong>Table 4.2.</strong> The coefficients of the polynomials and the $R^2$ for each DLR interval.</td>
<td>45</td>
</tr>
<tr>
<td><strong>Table 4.3.</strong> The coefficients of the polynomials for Partition A and B. Units as in Table 4.2.</td>
<td>45</td>
</tr>
<tr>
<td><strong>Table 4.4.</strong> Regression results of Models II-1, II-2 and II-3.</td>
<td>47</td>
</tr>
<tr>
<td><strong>Table 4.5.</strong> Description and regression results of Models V-1 to V-4 and VI-1 to VI-6 which estimate the cloudy-sky DLR from different cloud base heights ($P_{cb}$).</td>
<td>53</td>
</tr>
<tr>
<td><strong>Table 4.6.</strong> Regression results of the cloudy-sky DLR estimation Models VIII-1 to VIII-10.</td>
<td>58</td>
</tr>
<tr>
<td><strong>Table 5.1.</strong> Sensitivities of the cloudy-sky DLR due to the error in the surface pressure.</td>
<td>100</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 1.1</td>
<td>The annual mean global energy balance for the earth-atmosphere system.</td>
<td>6</td>
</tr>
<tr>
<td>Fig. 2.1</td>
<td>The spectrum of the downward LW specific intensities for clear-sky and overcast conditions with black cloud at different cloud base heights.</td>
<td>15</td>
</tr>
<tr>
<td>Fig. 2.2</td>
<td>Correlations of cloudy-sky DLR with the corresponding cloud base temperatures and the precipitable waters.</td>
<td>20</td>
</tr>
<tr>
<td>Fig. 3.1</td>
<td>Mean and standard deviation of the temperature profiles from the Phillips soundings.</td>
<td>31</td>
</tr>
<tr>
<td>Fig. 3.2</td>
<td>Mean and standard deviation of the water vapor profiles from the Phillips soundings.</td>
<td>32</td>
</tr>
<tr>
<td>Fig. 3.3</td>
<td>The McClatchey temperature and water vapor profiles.</td>
<td>33</td>
</tr>
<tr>
<td>Fig. 3.4</td>
<td>Difference between the air and the skin temperatures.</td>
<td>34</td>
</tr>
<tr>
<td>Fig. 4.1</td>
<td>An example of the fractional variance explained and residual regression error as variables are added in the stepwise regression.</td>
<td>61</td>
</tr>
<tr>
<td>Fig. 4.2</td>
<td>Residual plots of Model I-1. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.</td>
<td>62</td>
</tr>
<tr>
<td>Fig. 4.3</td>
<td>Residual plots of Model I-2.</td>
<td>63</td>
</tr>
<tr>
<td>Fig. 4.4</td>
<td>Residual plots of Model I-3.</td>
<td>64</td>
</tr>
<tr>
<td>Fig. 4.5</td>
<td>Residual plots of Model I-4.</td>
<td>65</td>
</tr>
<tr>
<td>Fig. 4.6</td>
<td>Residual plots of Model I-5.</td>
<td>66</td>
</tr>
<tr>
<td>Fig. 4.7</td>
<td>Emissivities defined by different brightness temperatures plotted against precipitable water.</td>
<td>67</td>
</tr>
</tbody>
</table>
Fig. 4.8. Emissivities defined by different brightness temperatures plotted against square root of the precipitable water. .........................................68

Fig. 4.9. Residual plots of Model II-1. ..............................................................69

Fig. 4.10. Residual plots of Model II-2. ...........................................................70

Fig. 4.11. Residual plots of Model II-3. ...........................................................71

Fig. 4.12. Residual plots of Model III-1. ........................................................72

Fig. 4.13. Residual plots of Model III-2a. .........................................................73

Fig. 4.14. Residual plots of Model III-2b. .........................................................74

Fig. 4.15. (a) Precipitable water vs. the ratio of brightness temperatures of channels 13 and 10; (b) Precipitable water vs. the ratio of brightness temperatures of channels 13 and 7. .................................................75

Fig. 4.16. Residual plots of Model IV-1. ..........................................................76

Fig. 4.17. Residual plots of Model IV-2. ..........................................................77

Fig. 4.18. Regression error and $R^2$ for DLR from clear and cloudy skies by linear models plotted with the corresponding cloud base pressure. ....................................................................................................78

Fig. 4.19. Statistics of cloud base temperature from 1584 Phillips soundings. (a) Mean $T_{cb}$; (b) Standard deviation of $T_{cb}$. ..............79

Fig. 4.20. Statistics of calculated DLR in partition A and B for clear and cloudy skies. (a) Mean DLR; (b) Standard deviation of DLR. .........80

Fig. 4.21. Statistics of calculated DLR for clear and cloudy skies.

(a) Mean DLR; (b) Standard deviation of DLR. ............................................81

Fig. 4.22. Residual plots of Model VI-2. ..........................................................82

Fig. 4.23. Transmissivity defined by Eq. (4.25) with the cloud base at 675 mb plotted as a function of the precipitable water. ...............83

Fig. 4.24. Residual plots of Model VII-1. ..........................................................84
Fig. 4.25. Residual plots of Model VIII-6. ...............................................................85

Fig. 4.26. Standard error of the resulting DLR and $R^2$ for the regression
models plotted with the corresponding cloud base pressure. .............86

Fig. 5.1. Sensitivities of the DLR to a 100% error in the cloud amount. ....101

Fig. 5.2. Sensitivities of the DLR due to the errors in the cloud base
heights for different average atmospheric conditions. .......................102

Fig. 5.3. Sensitivities of the clear-sky DLR due to the errors in the
surface pressure. ....................................................................................103

Fig. 5.4. Mean sensitivities of the DLR due to the errors in the surface
pressures in clear and cloudy-sky conditions with 10 mb
uncertainties in the surface pressures. ..............................................104

Fig. 5.5. RMS errors for regression model IV-1 for clear-sky conditions
with nominal instrument noise included. .............................................105

Fig. 5.6. As in Fig. 5.5, but for average and standard deviation. .............106

Fig. 5.7. As in Fig. 5.6 but for model I-1. ......................................................107

Fig. 5.8. RMS errors for clear and cloudy-sky conditions with and
without nominal instrument noises. .......................................................108

Fig. 5.9. DLR errors due to errors from uncertainties in cloud amounts
and cloud base heights. .........................................................................109

Fig. 5.10. Combined DLR errors contributed from regression error, and
the errors from the uncertainties in cloud amounts and cloud base
heights. Curves are shown for the indicated cloud amounts.
Different degrees of uncertainties in cloud parameters are
considered: (a) $\delta A=\pm 10\%$ and $\delta P_{cb}=\pm 50$ mb, (b) $\delta A=\pm 10\%$ and
$\delta P_{cb}=\pm 100$ mb. ................................................................................110
Fig. 5.11. As in Fig. 5.10 but for (a) $\delta A = \pm 30\%$ and $\delta Pcb = \pm 50$ mb, (b) $\delta A = \pm 30\%$ and $\delta Pcb = \pm 100$ mb.

Fig. 6.1. Skin Temperature (°K) from the NMC analysis at 00Z, 12/22/90. Contours are at 10°C intervals.

Fig. 6.2. Column precipitable water (cm) from the NMC analysis at 00Z, 12/22/90.

Fig. 6.3. Clear-sky DLR (Wm$^{-2}$) calculated by radiative transfer model with the NMC analysis at 00Z, 12/22/90.

Fig. 6.4. Zonal statistics of the calculated (solid) and estimated (dashed and dotted) DLR (Wm$^{-2}$) at 00Z, 12/22/90.

Fig. 6.5. Zonal statistics of the surface temperature (°K) from the NMC analysis at 00Z, 12/22/90.

Fig. 6.6. Zonal mean (solid) and standard deviation (dotted) of the column precipitable water (cm) from the NMC analysis at 00Z, 12/22/90.

Fig. 6.7. Zonal mean (solid) and standard deviation (dotted) of the topography (km) where sea surface height is assumed at 0.

Fig. 6.8. Clear-sky DLR (Wm$^{-2}$) estimated by regression model IV-1 with simulated HIRS radiances and the NMC surface pressure at 00Z, 12/22/90.

Fig. 6.9. As in Fig. 6.8 but with the estimated surface pressure.

Fig. 6.10. Flux difference field (Wm$^{-2}$) between the estimated DLR (with the NMC surface pressure) and calculated DLR for 00Z, 12/22/90.

Fig. 6.11. As in Fig. 6.10 but with the estimated surface pressure.
Fig. 6.12. Histogram for the percentage of the total grid points falling into the given ranges of the DLR differences for 00Z, 12/22/90.

Fig. 6.13. Zonal mean estimated minus calculated DLR for 00Z, 12/22/90. The differences associated with the NMC and estimated Ps are plotted by solid and dashed curves, respectively.

Fig. 6.14. DLR estimated using the NMC surface pressure versus the calculated DLR for 00Z, 12/22/90.

Fig. 6.15. As in Fig. 6.14 but with estimated surface pressure.

Fig. 6.16. Surface pressure from the NMC analysis on 00Z, 12/22/90.

Fig. 6.17. (a) DLR difference versus surface temperature, and (b) DLR difference versus precipitable water for 00Z, 12/22/90. DLR is estimated with the NMC surface pressure.

Fig. 6.18. As in Fig. 6.17 but for the estimated surface pressures.

Fig. 6.19. DLR difference versus the NMC surface pressure for 00Z, 12/22/90.

Fig. 6.20. As in Fig. 6.19 but for the estimated surface pressures.
CHAPTER I
Introduction

Considering the earth-atmosphere as a closed system, the solar energy is the only external energy system source. Through millions of years, this system has maintained its total energy system in approximately a constant state. This energy balance is achieved mainly by the radiative processes.

At the top of the atmosphere, the incoming solar radiation is counterbalanced by the reflected solar radiation and thermal radiation emitted by the atmosphere and the earth's surface. While in the atmosphere and at the earth's surface, the energy balance is accomplished by the radiative processes as well as by the latent heat and sensible heat exchanges (see Fig. 1.1).

The energy balance of the earth-atmosphere system determines a mean terrestrial temperature for the system while the latitudinal energy distribution largely determines the climate. All the relevant processes must be studied in order to understand the way the energy balance is achieved and also to understand the global and regional climates which are heavily influenced by the energy budget.

The surface energy budget plays an important role in determining many atmospheric/oceanic processes, especially on the global scale. One needs a good estimation of surface energy budget to explain a large scale phenomenon like El Niño/Southern Oscillation (ENSO) (see TOGA Scientific Plan, 1985). The accuracy required for the surface energy budget
estimation is considered to be about 10 Wm\(^{-2}\) which corresponds to a 1\(^{\circ}\)K error in the ocean temperature (TOGA, 1985). Many other scientific programs like the Global Energy and Water Cycle Experiment (GEWEX) (WCRP-5, 1988), and the World Climate Research Program (WCRP, 1987) have shown the need for a better estimation of the surface energy budget.

The downward terrestrial radiation (or known as downward longwave radiation, DLR) at the earth's surface is one of the necessary components in the study of surface energy budget. It is very difficult to study the global distribution of DLR with the use of ground observations because observations are sparse or non-existent, particularly over the oceans, and the available data are of questionable accuracy. Earth-orbiting satellites can provide global coverage with high spatial resolution and hence we want to develop a scheme to obtain the DLR from satellite's observations.

It has been shown that satellites are the most powerful tools for observing the radiation at the top of the atmosphere, both incoming or outgoing (see Barkstrom and Smith, 1986; ERBE, 1981). However, global analysis of surface radiation budget based on satellite observations is still not well-developed, especially for the DLR. A comprehensive review of the methods using satellite observations for studying the surface radiation budget has been done by Schmetz (1989). For more than a decade, the surface solar radiation budget based on satellite measurements has been widely studied, e.g., Tarpley, 1979; Gautier \textit{et al.}, 1980; Pinker and Corio, 1984; Dedieu \textit{et al.}, 1986; etc. (see Schmetz, 1989). In the longwave range, efforts have been made by Smith and Woolf (1983), McMillin and Govindaraju (1983), Gupta \textit{et al.} (1983), Meerkcoetter and Grassl (1984),

These methods may be classified into two categories: statistical and physical. Statistical methods usually use a detailed radiative transfer model with in-situ soundings to synthesize the outgoing radiances for a particular instrument and to calculate the DLR or the net longwave flux at the surface. Regression analysis is performed for the DLR and the 'independent variables' – usually the radiances from different channels of a satellite instrument. Physical methods use either a detailed or a parameterized radiative transfer model with quantities inferred from satellite observations, such as the temperature profile, layer water vapor content, cloud fraction and cloud top height. Analyzed field and empirical bulk formulae are used in some methods as well.

Recently, the WCRP has begun experimental processing of the shortwave surface radiation budget using techniques developed by Prof. Pinker and her colleagues. However, none of the longwave techniques have been found to be acceptable for routine analysis (R. G. Ellingson, personal communication).

Schmetz et al. (1986) adapted empirical formulae to estimate the DLR as a function of temperatures at surface, 1000 mb and 850 mb, surface pressure, cloud amount, solar cloud transmittance, cloud top height and mean topographic altitude. The thermodynamic data are obtained from analyzed synoptic grid data and the cloud parameters are derived from observations by METEOSAT-2. The estimation is limited to the area with ground synoptic observations and is limited to daytime use only because
the visible channel is needed to determine the solar cloud transmittance, which is used to empirically determine the cloud base height and cloud emittance.

The studies by Darnell et al. (1983) and Frouin and Gautier (1988) use a physical method. The atmospheric data necessary for the model, profiles of temperature, water vapor and ozone as well as surface temperature and cloud amount, are estimated from satellite radiance measurements. The drawbacks of this kind of scheme are that the technique requires a large amount of computer time necessary for a global application and errors in retrieving temperature and water vapor profiles propagate into the flux calculations.

Darnell et al. (1990) studied the surface radiation budget using the C1 data from the International Satellite Cloud Climatology Project (ISCCP) (Schiffer and Rossow, 1985). The DLR component was estimated by a parameterized model of Gupta (1989) adapted to operate with the C1 data. Although the parameterized model does not need intensive calculations, since it uses the satellite retrieved temperature and moisture profiles as input data, the accuracy of the estimation may be limited to that of the retrievals due to error propagation. Their study showed that the majority of the error in the flux estimation is believed to be caused by the uncertainties in the input data.

In view of the difficulties discussed above and because of the importance of the surface longwave radiation budget, it appeared that a new approach to the longwave problem was necessary. The purpose of this research is to develop a new technique of estimating the DLR which overcomes many of the shortcomings of previous approaches. The new
technique was designed to use radiance observations from the NOAA operational satellites. In particular, instead of using derived temperature and water vapor profiles in a radiation model, this technique estimates the flux directly from radiance using statistical information obtained from a large set of radiation model calculations based on observed data.

In the following material, Chapter II describes the basic equations governing the DLR and the conceptual framework of the DLR estimation technique; Chapter III describes the data sources for both the radiative transfer model and the regression; Chapter IV discusses the model development and describes various regression models; Chapter V discusses the sensitivity of the DLR to the various error components; Chapter VI presents the results of the validation; and Chapter VII summarizes the DLR estimation technique and discusses the future studies for this technique.
CHAPTER II
Theoretical Basis

Basic Equations

The upwelling radiance observed at a satellite is related to the specific intensity \( I_v \) at a given local zenith angle \( \theta \) as

\[
N_i(\mu) = \int I_v(z_t, \mu) \phi_i(v) dv
\]  \hspace{1cm} (2.1)

where \( \mu \equiv \cos \theta \), \( \phi_i \) is the normalized response function for channel \( i \), \( v \) is the wave number or frequency, and \( z_t \) is the satellite altitude or the top of the atmosphere.

The upwelling clear column radiance for a plane parallel, horizontally homogeneous non-scattering atmosphere in local thermodynamic equilibrium above a black surface is related to the temperature and absorbing constituents as

\[
I_v(z_t; \mu) = B_v'(0) T_v(z_t, 0; \mu) + \int_0^{z_t} B_v(z') \frac{\partial T_v(z_t, z'; \mu)}{\partial z'} dz'
\]  \hspace{1cm} (2.2)

where \( B_v'(0) \) is the Planck function evaluated with the surface temperature, \( B_v(z') \) is the Planck function evaluated with the temperature at level \( z' \), and the monochromatic transmittance \( T_v(z_t, z'; \mu) \equiv e^{-(\tau_v(z) - \tau_v(z'))/\mu} \), where \( \tau_v \) is the optical depth for the atmospheric absorbers.

Since the liquid water clouds are almost opaque in the thermal infrared (IR), the transmissivity through the cloud is nearly zero if the cloud thickness exceeds about 300 meters. Therefore the upwelling specific intensity for a cloudy scene may be approximated as
\[ I_{\nu,c}(z_t, \mu) = B_{\nu}(z_{ct}) T_{\nu}(z_t, z_{ct}; \mu) + \int_{z_{ct}}^{z_t} B_{\nu}(z') \frac{\partial T_{\nu}(z_t, z'; \mu)}{\partial z'} dz' \] (2.3)

where \( z_{ct} \) is the cloud top height and the monochromatic transmittance is defined as above except \( \tau_{\nu} \) is the optical depth for non-cloud absorbers.

The DLR is related to the downward specific intensity \( I_{\nu}(z = 0, -\mu) \) at the earth’s surface in an axial-symmetric atmosphere as

\[ DLR = 2\pi \int_0^\infty \int_\mu I_{\nu}(z = 0, -\mu) \mu d\mu d\nu \] (2.4)

For clear sky conditions we can write

\[ I_{\nu,0}(z = 0; -\mu) = - \int_0^{z_t} B_{\nu}(z') \frac{\partial T_{\nu}(0, z'; -\mu)}{\partial z'} dz' \] (2.5)

For overcast, black cloud conditions we can write

\[ I_{\nu,c}(z = 0; -\mu) = B_{\nu}(z_{cb}) T_{\nu}(0, z_{cb}; -\mu) - \int_0^{z_{cb}} B_{\nu}(z') \frac{\partial T_{\nu}(0, z'; -\mu)}{\partial z'} dz' \] (2.6)

where the \( z_{cb} \) is the cloud base height.

For partial cloud conditions, assuming an infinitesimally thin, single layer cloud with amount \( A \), the total DLR is commonly written as the cloud-amount-weighted sum of the DLR from clear and cloudy portions of the sky, respectively, as

\[ DLR = (1 - A) F_{o}^{\downarrow} + A F_{c}^{\downarrow} \] (2.7)

where \( F_{o}^{\downarrow} \) and \( F_{c}^{\downarrow} \) are the downward longwave fluxes at the surface from the clear and cloudy portions of sky, respectively. For none-black clouds, \( A \) is the effective fractional cloud amount which is defined as the product of the cloud emissivity and the absolute fractional cloud amount.
The brightness temperature $T_b$ corresponding to a satellite observed radiance $N_i$ is defined as the temperature at which a black body emits the same amount of radiation seen by the radiometer. That is,

$$N_i = \int_{\Delta \nu_i} B_{\nu}(T_b) \phi_i(\nu) d\nu$$  \hspace{1cm} (2.8)

A conversion formula from NOAA Polar Orbiter Users Guide (NOAA, 1984) could calculate the brightness temperature directly with two band correction coefficients, $b$ and $c$, for each of the HIRS channels.

$$T_b = \frac{T^* - b}{c}$$  \hspace{1cm} (2.9)

where $T^* = B_{\nu_o}^i(N)$ is the apparent brightness temperature that is the inverse of Planck’s function calculated with the central frequency $\nu_o$ of the channel and the corresponding radiance $N$.

**Conceptual Framework**

Theoretically one can calculate the DLR using a radiative transfer model with given temperature and atmospheric gases profiles. From satellite observations, existing temperature and water vapor retrieval techniques can provide the needed information for such models to proceed to global DLR studies. Highly parameterized radiative transfer models designed to use satellite-inferred information have been well developed (Darnell et al., 1983; Gupta et al., 1983; Gupta, 1989; Frouin and Gautier, 1988; Schmetz, 1984; and Schmetz et al., 1986). We call them ‘physical’ methods in the sense that the calculations employ parameters derived from basic physics. In some cases, analyzed fields from synoptic observations or numerical weather prediction models are also used to provide extra
information or to fill in the missing data whenever retrievals are not available.

For quite a long time, oceanographers and meteorologists in the study of air-sea interaction have used empirically-derived ‘bulk formulae’ to estimate net LW radiation based on synoptically observed variables (see Fung et al., 1984). Darnell et al. have parameterized Gupta’s LW model (Gupta, 1989) by regression to use specifically with ISCCP C1 data set, which includes satellite-inferred temperature, water vapor and cloud information, to estimate the DLR (Darnell et al., 1990). We call these types of approaches ‘statistical’ techniques.

Our approach to the DLR estimation is also ‘statistical’. However, we explored the possibility of using satellite observations directly instead of their inferred quantities. Furthermore, we designed the technique for possible operational implementation so that a long-term global DLR field will be available in the future.

Different from all the existing methods that use inferred information from satellite observations, using satellite observations directly is possible to reduce the error of DLR estimation that propagates from retrievals to either physical or statistical models. Furthermore, the computational speed of statistical models is superior compared to physical models. This feature is important for operational use and for using historical data to derive the radiation budget climatology.

Since the physical and statistical techniques discussed above use inferred temperature and moisture profiles, the errors in the inference schemes propagate into the derived radiation products. If the radiance
observations would be used directly in the DLR estimation scheme, it would be possible to remove this source of errors.

As described in (1.1) and (1.2), satellite-born radiometers measure the upwelling radiance emerging from the earth-atmosphere system. When clouds are present in the field of view of the satellite-borne radiometer, the clouds contaminate the information about the temperature and absorbing gas concentration distributions. In addition, when clouds are optically thick there is little information concerning cloud base and the atmospheric conditions underneath (sometimes called decoupling; Schmetz, 1989). The DLR contributed by the clouds can not be well estimated by the cloud contaminated upwelling radiance unless information below the cloud base is provided by other means.

An alternative is to use cloud-cleared radiance. Clear-column radiance, which is inferrable and is operationally derived from radiance observations (Smith and Woolf, 1976; McMillin, 1978; McMillin and Dean, 1982), always contains the information for the whole atmospheric column even in a cloudy sky. Such data are used to infer temperature and moisture profiles. Particularly, it consistently provides the information for the near-surface temperature and water vapor fields that are the most important parameters for estimating the DLR. However, the cloud-clearing process can be only carried out in a partly cloudy scene. That means whenever a unit area is under overcast condition, the clear-column radiance will not be available. The size of the unit area changes depending on the size of instrument’s field of view and the cloud-clearing techniques. For HIRS/2, it ranges from 50x50 km² to 180x90 km².
For any single day, the clear-column radiance probably would never have a 100% global coverage. A HIRS clear-column radiance data set for the period from Dec. 15, 1990 to Jan. 14, 1992 gives an about 82% of daily global coverage with both orbits. Accumulating the data for 14 days, the coverage rate rises to 91% and it reaches 97% for 32 days. Except for the geographical regions where it is consistently overcast, it is legitimate to say that the global coverage of clear-column radiance is sufficient for the study of monthly mean DLR. If the scheme works, it might be possible to get cloud-cleared radiances from a global analysis using the objective analysis schemes employed by numerical weather prediction models.

A technique that uses linear combination of the satellite observed radiances to estimate the outgoing longwave radiation (OLR) has been proven successful (see Ellingson et al., 1988, 1993). For the OLR estimation, the satellite observed radiances are portions of the total OLR spectrum and the linear combination of the radiances can be interpreted as a weighted sum for the total OLR.

The DLR estimation using radiances at the top of the atmosphere is not physically as straightforward as the OLR estimation. Although it is possible to estimate the DLR from a linear combination of clear-column radiances, as will be shown, the radiances are not part of the DLR. The upwelling radiances only reflect characteristics of the temperature and moisture fields that influence the DLR. It is a kind of empirical transformation from radiance to the needed information. Finding the appropriate functional forms is the major task of this study.

Mathematically, one can not find an analytical expression for the DLR in terms of the clear-column radiances except for assumed dependence of
the atmospheric structure (e.g., isothermal condition or the thermal emission is linear in the optical depth) because they are all complicated functionals. However, physically we do know that the clear-column radiances contain information related to the DLR (i.e., the information of the vertical structures of temperature and water vapor). To find out what are the most important parameters to the DLR, we need to look at the spectrum of downward longwave specific intensity first.

Fig. 2.1 shows the spectrum of the downward LW specific intensities calculated by the Ellingson and Serafinao (1984) longwave radiative transfer model with observed temperature and water vapor profiles for clear-sky and overcast conditions with black clouds at different cloud base heights. One can easily perceive that there are several spectral intervals where the intensities do not have noticeable variation from clear-sky to any cloudy-sky conditions. Adding the clouds or changing the atmospheric conditions aloft do not make any significant difference to intensities in those intervals. Those are the spectral intervals where strong absorptions from atmospheric gases occur. They are the $H_2O$ pure rotational band (0-600 cm$^{-1}$), the 6.3 μm $H_2O$ band, and the 15 μm and 4.3 μm $CO_2$ band systems. In the 9.6 μm $O_3$ absorption band, the downward specific intensity is a combination of the weak $H_2O$ continuum emission and the strong $O_3$ emission attenuated by the $H_2O$ absorption. Owing to the vertical distribution of $O_3$, though the 9.6 μm $O_3$ absorption band is strong, the downward specific intensity from this spectral interval represents an effective temperature for the stratosphere where most of the $O_3$ is located. Except for the 9.6 μm $O_3$ band, the effective brightness temperatures of the downward specific intensities from the strong absorption spectral intervals
are very close to the surface temperatures because the concentration of $CO_2$ and $H_2O$ in the lowest layer of the atmosphere is sufficient to absorb and emit almost completely. In other words, they are optically ‘near-opaque’. However, this does not apply to relatively dry air where the water vapor is not abundant enough for saturating the absorption, particularly in the pure rotational $H_2O$ band, although the absorption is relatively strong. This occurs mostly in the polar and desert regions. On average, one could estimate the flux from these spectral intervals quite accurately with only the near-surface temperature.

On the other hand, there are some spectral intervals where absorption by atmospheric gases is relatively weak or optically ‘near-transparent’. For example, the atmospheric window (about 800-1250 cm$^{-1}$ excluding the 9.6 $\mu$m O$_3$ band) where weak $H_2O$ continuum absorption is the determining factor. There are some other ‘windows’ in between of the strong absorption band systems. They are not optically as transparent as the 8-12 $\mu$m window, and they are in a mixture of the water vapor continuum and the wings of the strong $H_2O$ and $CO_2$ absorption bands. Due to the optical transparency, the downward specific intensity will then be determined integrally by the vertical distributions of the atmospheric gases, mainly $H_2O$, and the temperature.
Because of the distinct characteristics of different portions of the DLR spectrum, we shall discuss and treat them separately. For convenience, we define Partition A to include the optically ‘near-opaque’ spectral intervals and Partition B to include the remainder of the spectrum, i.e., the optically ‘near-transparent’ spectral intervals. Fig. 2.1 also shows the partitions of the DLR spectrum. The ranges of the ‘near-opaque’ spectral intervals are a little smaller than the ones usually defined for those band systems. We exclude the wings of the band systems so that intervals in Partition A are as

![Fig. 2.1. The spectrum of the downward LW specific intensities for clear-sky and overcast-sky with black cloud at different cloud base heights. The shaded portions are the ‘near-opaque’ spectral intervals, namely, Partition A. The rest are the ‘near-transparent’ spectral intervals, the Partition B.](image)
'opaque' as possible. Specifically, Partition A includes the intervals of 0-400, 600-700, 1320-1800 and 2240-2400 cm⁻¹. Partition B includes the intervals of 400-600, 700-1320, 1800-2240 and 2400-3000 cm⁻¹.

From (2.4), the monochromatic downward LW flux from a clear sky can be expressed as

\[
F_{ν,↓} = -\int_{0}^{z_t} \pi B_{ν}(z′) \frac{∂T_{Fν}(0, z′)}{∂z′} dz′
\]

(2.10)

where \( T_{Fν}(0, z′) \equiv 2\int_{0}^{z′} T_{ν}(0, z′; μ) μ dμ \) is the monochromatic flux transmittance. One can always find an effective temperature so that

\[
F_{ν,↓} = \pi B_{ν}(0, z_t) \varepsilon_{Fν}(0, z_t)
\]

(2.11)

where \( \varepsilon_{Fν} = 1 - T_{Fν} \) is the monochromatic flux emittance, and the effective temperature is defined by

\[
\pi B_{ν}^{(0, z_t)} ≡ \frac{-\int_{0}^{z_t} \pi B_{ν}(z′) \frac{∂T_{Fν}(0, z′)}{∂z′} dz′}{\varepsilon_{Fν}(0, z_t)}
\]

(2.12)

Similar expression for a cloudy sky with a single layer of overcast black cloud at an altitude \( z_c \) is

\[
F_{ν,c} = \pi B_{ν}(z_c) T_{Fν}(0, z_c) + \pi B_{ν}^{(0, z_c)} \varepsilon_{Fν}(0, z_c)
\]

(2.13a)

or

\[
F_{ν,c} = \pi B_{ν}(z_c) + \varepsilon_{Fν}(0, z_c) (\pi B_{ν}^{(0, z_c)} - \pi B_{ν}(z_c))
\]

(2.13b)

\( \varepsilon_{Fν}(0, z′) \) can be considered as the emissivity of the atmosphere in the layer between surface and \( z′ \). \( B_{ν}^{(0, z_c)} \) is defined similar to (2.12) by replacing \( z_t \) to \( z_c \). It represents an effective emitting temperature for the layer between surface and the cloud base.
Generally we could assume the transmittance function as an exponential decay with the power of the optical path. Hence the emittance is approximately proportional to the optical path for a small optical path and asymptotically equals to one for a very large optical path. Let’s consider the extreme cases. First, in the very strong absorption bands, the fluxes from those spectral intervals would be just a function of the effective emitting temperature which corresponds to the $\bar{B}_\nu$’s in (2.11) and (2.13a). Fig. 2.1 suggests that both $\bar{B}_\nu$’s are very close to the surface temperature. For the other extreme where absorption is weak, the clear sky DLR varies linearly with the optical path with a fixed temperature profile. The DLR from a cloudy sky would be determined predominantly by $B_\nu(z_c)$ if either the emittance is small or the difference between $\bar{B}_\nu(0,z_c)$ and $B_\nu(z_c)$ is small (see (2.13b)). This applies to a dry atmosphere or a low cloud situation.

Thus, in most cases, we only need to know the near-surface temperature to estimate the DLR in Partition A. However, absorber concentrations in addition to the temperature field are needed for estimating DLR in Partition B. Among all absorbers, water vapor is the most important one due to both its spectral positions and variability in concentration.

In order to gain physical insights, we parameterize the transmittance in terms of the precipitable water. In the spectral intervals where water vapor continuum absorption dominates, the absorption coefficient is linearly proportional to atmospheric pressure, $P$, and water vapor partial pressure, $e$,

$$k_\nu = k_{\nu 1} \frac{P}{P_o} + k_{\nu 2} \frac{e}{P_o}$$  \hspace{1cm} (2.14)
where \( P_o \) is the pressure of one atmosphere. The optical path \( \tau_V \) is

\[
\tau_V = \int_0^{z_i} k_V \rho_v dz
\]

(2.15)

Assuming that the water vapor mixing ratio \( q \) is distributed as

\[
q = q_o \left( \frac{P}{P_o} \right)^{\lambda}
\]

(2.16)

where \( q_o \) is the water vapor density at pressure \( P_o \). With the definition of precipitable water \( PW \),

\[
PW = \int_0^{z_i} \rho_v dz
\]

(2.17)

where \( \rho_l \) is the density of liquid water, one can then express \( \tau_V \) in terms of precipitable water \( PW \)

\[
\tau_V = k_{v1}^* \cdot PW + k_{v2}^* \cdot PW^2
\]

(2.18)

where \( k_{v1}^* = \frac{k_v \rho_l (\lambda + 1)}{(\lambda + 2)} \), \( k_{v2}^* = \frac{k_v \rho_l^2 g (\lambda + 1)}{2 \cdot \epsilon} \), \( \epsilon = 0.622 \) and \( g \) is gravity. The transmittance from surface to space due to water vapor alone can be expressed as a function of \( PW \)

\[
T_v = \text{Exp}\{- (k_{v1}^* \cdot PW + k_{v2}^* \cdot PW^2)\}
\]

(2.19)

and we can write the emittance for small optical paths approximately as

\[
\mathcal{E}_V \equiv k_{v1}^* \cdot PW + k_{v2}^* \cdot PW^2
\]

(2.20)

This gives a clue that in the spectral intervals where the water vapor continuum absorption dominates, we can expect the fluxes from these portions of DLR spectrum to be related to precipitable water in a similar manner seen in (2.11) and (2.20). This illustrates the role of water vapor in modulating the DLR. It is necessary to estimate the amount of water vapor
(or precipitable water), implicitly or explicitly, in our technique so that the
DLR will be estimated correctly.

For cloudy conditions, the cloud base temperature is usually more
crucial than water vapor amount in determining the DLR in Partition B
particularly for low and middle clouds (see Fig. 2.2). The cloud base height
or cloud base temperature currently cannot be retrieved through satellite
observations. However, it can be estimated from several sources. With the
radiance derived effective cloud top, which is one of the operational TOVS
products, one can subtract a climatological cloud thickness to obtain the
cloud base position (Chou, 1989). The International Satellite Cloud
Climatology Project (ISCCP, see WMO, 1988) data set reports cloud base
and then a cloud base temperature can be obtained with the retrieved
temperature profile (Darnell et al., 1990). The Air Force’s Real Time
Nephanalysis (RTNeph) operationally reports the cloud amounts, cloud top
and base heights for up to four layers of clouds (AFGWC, 1988). The
RTNeph is considered to be the best cloud data source for DLR estimation
purpose, because not only that its cloud data are obtained from all possible
observations and satellite retrievals, but also that, in particular, it provides
cloud data for a multi-layer cloud structure. Synoptical cloud data from
long-term ground observations (Hahn et al., 1982, 1984; Warren et al., 1986,
1988) provides a cloud climatology and can be used when there is no other
better cloud information.
In the estimation of cloudy sky DLR, uncertainties of both cloud base height and cloud amount induce errors. A technique is considered better if it is more insensitive to these two quantities.

Since the technique is aimed for application to operational global DLR estimation, we have chosen to design it to the second-generation, high resolution infrared spectrometer (HIRS/2) on board of the NOAA TIROS-N series satellites (see NOAA Polar Orbiter Data User Guide, 1984). Table 2.1 shows the specifications and typical characteristics of the HIRS/2 channels. The first seven channels are in the 15 µm CO$_2$ band and are usually used to determine temperature profile, channels 8 and 9 are window channels (although channel 9 is located in the 9.6 µm O$_3$ band), and channels 10 to 12, located in the 6.3 µm H$_2$O band, are used to estimate water vapor distribution.

![Correlation Coefficient](image)

**Fig. 2.2.** Correlations of cloudy-sky DLR with the corresponding cloud base temperatures and the precipitable waters.
As in many studies where appropriate observations are not available, we have to simulate the radiation data by a radiative transfer model. We calculated the DLR, the clear-column radiances and the corresponding brightness temperatures from 1600 soundings using the radiative transfer model originally developed by Ellingson and Gille (1978). The model and the calculations will be described in detail in the next chapter.

From the conceptual discussion and the theoretical derivations, we concluded two general models that will be used to estimate the DLR. First is the Linear models:

\[
F_o^\downarrow = a_o + \sum a_if_i(N) \tag{2.21}
\]

\[
F_c^\downarrow = b_o(z_{cb}) + \sum b_if_i(z_{cb})f_i(N) \tag{2.22}
\]

where \(f_i\) are functions of radiances. They are either linear or non-linear in radiances, e.g., cross-products of radiances from different channels. In an overcast sky, the coefficients \(b\)'s are functions of cloud base height.

The second is the Emissivity-approach models:

\[
DLR = \overline{B} \cdot \mathcal{E}(pw) \tag{2.23}
\]

where \(\overline{B}\) represents an effective emitting temperature and \(\mathcal{E}(PW)\) stands an effective atmospheric emissivity in terms of precipitable water. Ultimately, precipitable water will be represented by radiances so that precipitable water is included implicitly and its physical dependence is kept correctly.

Due to the different dominating physical mechanisms, each interval of the DLR spectrum may behave differently in a statistical sense. Regression analyses for each model are studied for different combination of
DLR spectral intervals. There are three possible cases: DLR from whole spectrum, two partitions each with four spectral intervals, and individual spectral intervals (see Fig. 2.1).
Table 2.1. Specifications and typical characteristics of HIRS/2 on NOAA-10 satellite (NOAA Polar Orbiter Data Users Guide, 1984)

<table>
<thead>
<tr>
<th>Channel number</th>
<th>Central wave number (cm⁻¹)</th>
<th>Description</th>
<th>Peak of $dT/dlnP$ (mb)</th>
<th>NEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>667.70</td>
<td>15 μm CO₂ Band</td>
<td>30</td>
<td>3.00</td>
</tr>
<tr>
<td>2</td>
<td>680.23</td>
<td>15 μm CO₂ Band</td>
<td>60</td>
<td>0.67</td>
</tr>
<tr>
<td>3</td>
<td>691.15</td>
<td>15 μm CO₂ Band</td>
<td>100</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>704.33</td>
<td>15 μm CO₂ Band</td>
<td>280</td>
<td>0.31</td>
</tr>
<tr>
<td>5</td>
<td>716.30</td>
<td>15 μm CO₂ Band</td>
<td>475</td>
<td>0.21</td>
</tr>
<tr>
<td>6</td>
<td>733.13</td>
<td>15 μm CO₂ Band</td>
<td>725</td>
<td>0.24</td>
</tr>
<tr>
<td>7</td>
<td>750.72</td>
<td>15 μm CO₂ Band</td>
<td>Surface</td>
<td>0.20</td>
</tr>
<tr>
<td>8</td>
<td>899.50</td>
<td>Window</td>
<td>Window</td>
<td>0.10</td>
</tr>
<tr>
<td>9</td>
<td>1029.01</td>
<td>9.3 μm O₃ Band</td>
<td>Window-O₃</td>
<td>0.15</td>
</tr>
<tr>
<td>10</td>
<td>1224.07</td>
<td>6.3 μm H₂O Band</td>
<td>Lower-Tropo. H₂O</td>
<td>0.16</td>
</tr>
<tr>
<td>11</td>
<td>1363.32</td>
<td>6.3 μm H₂O Band</td>
<td>Mid-Tropo. H₂O</td>
<td>0.20</td>
</tr>
<tr>
<td>12</td>
<td>1489.42</td>
<td>6.3 μm H₂O Band</td>
<td>Upper-Tropo. H₂O</td>
<td>0.19</td>
</tr>
<tr>
<td>13</td>
<td>2191.38</td>
<td>4.3 μm CO₂ Band</td>
<td>Surface</td>
<td>0.006</td>
</tr>
<tr>
<td>14</td>
<td>2208.74</td>
<td>4.3 μm CO₂ Band</td>
<td>650</td>
<td>0.003</td>
</tr>
<tr>
<td>15</td>
<td>2237.49</td>
<td>4.3 μm CO₂ Band</td>
<td>340</td>
<td>0.004</td>
</tr>
<tr>
<td>16</td>
<td>2269.09</td>
<td>4.3 μm CO₂ Band</td>
<td>170</td>
<td>0.002</td>
</tr>
<tr>
<td>17</td>
<td>2360.00</td>
<td>4.3 μm CO₂ Band</td>
<td>15</td>
<td>0.002</td>
</tr>
<tr>
<td>18</td>
<td>2514.58</td>
<td>4.3 μm CO₂ Band</td>
<td>Window-Solar</td>
<td>0.002</td>
</tr>
<tr>
<td>19</td>
<td>2665.38</td>
<td>Window</td>
<td>Window-Solar</td>
<td>0.001</td>
</tr>
</tbody>
</table>

- NEAN is the noise equivalent differential radiance and is in unit of milliwatts m⁻² sr⁻¹ / cm⁻¹
CHAPTER III

Data for Regression Analyses and Validation

Representativeness of the sample is essential for establishing the relationships for physical variables using statistical methods. While large amounts of DLR observations are not available over space and time, the model simulation is one of the alternatives. Sounding data are observed operationally so that it is relatively easier to compile a sounding data set representing the global and seasonal distributions. A radiative transfer model can simulate the DLR and the satellite radiance observations using this sounding data set as the input. It is assumed that the radiation data generated herein will be as representative as the soundings.

While it is again for lack of the observations to validate the developed DLR estimation technique, the best we can do is to simulate the DLR fields using soundings from the independent sources. It would be better to use the satellite observed radiances for the validation, however, due to the availability of the radiance data, we cannot use the real clear-sky radiance observations but have to use the simulated ones.

In the following sections, the data sources used in the regression analyses and the validation will be described. The radiative transfer model which was used to calculate the DLR and the upwelling clear-sky radiances will be first discussed.
Radiative Transfer Model

The approach followed herein is to simulate the DLR and upwelling clear-sky radiances, as needed to determine the estimation equations statistically, from a radiative transfer model. In particular, this study uses the model first reported by Ellingson and Gille (1978) and updated by Ellingson and Serafino (1984) and Ellingson et al. (1992). This model has been tested in the project of the Intercomparison of Radiation Codes for Climate Models (ICRCCM, see Ellingson et al., 1991) and with observations. Although the absolute accuracy has not been established, the model has been shown to produce the OLR (Outgoing Longwave Radiation) and DLR calculations that agree with the line-by-line calculations to within about ±2%. Furthermore, Ellingson and Sarefino (1984) showed agreement with aircraft OLR and DLR data to within the accuracy of the observations.

This radiative transfer model has 140 spectral intervals across 0-3000 cm\(^{-1}\) with interval widths varying from 5 cm\(^{-1}\) to 40 cm\(^{-1}\) depending on their locations in the spectrum. There are five gases included in the model: water vapor (H\(_2\)O), ozone (O\(_3\)), carbon dioxide (CO\(_2\)), nitrous oxide (N\(_2\)O) and methane (CH\(_4\)). Carbon dioxide concentration is assumed constant at 330 ppm by volume. Concentrations of N\(_2\)O and CH\(_4\) are also assumed constant at 0.279 ppm and 1.75 ppm, respectively.

The Malkmus random band model (Malkmus, 1967) was used for the local lines of all of the gases. A mixture of models were used to estimate water vapor continuum absorption (Ellingson et al., 1989). The effect of overlapping absorptions from different gases is calculated as the product of the individual transmissivities.
For each sounding, the specific intensities for all model spectral intervals were calculated at the earth’s surface and the top of the atmosphere for clear and cloudy conditions. In cloudy conditions, black clouds are added at different pressure levels with the temperature and water vapor profiles unchanged.

Using the NOAA supplied, laboratory measured values of HIRS/2 response functions (or filter functions), the nadir upwelling clear-column radiances at the top of the atmosphere for all HIRS/2 channels were simulated by trapezoidal integration of (2.1). The filter functions were first averaged to the corresponding model spectral resolution. As a result of this, the calculated radiances may be different from the radiances observed by the radiometer. In order to account for this discrepancy, the following procedure was devised.

The mapping tables were first generated for all HIRS/2 channels between the model brightness temperatures and the model calculated radiances by the trapezoidal integration mentioned herein before using the same spectral-interval-averaged filter functions but with the Planck function instead of the model-calculated intensities (i.e., Eq. (2.8)). With these mapping tables, one can find the corresponding brightness temperatures for the given model calculated radiance values. Using the NOAA supplied formula which accounts for the band effect of the channels, the ‘corrected radiances’ were calculated from the Planck function with the apparent temperatures $T^*$ (see Eq. (2.9)). The deviation between the model radiances and the NOAA radiances is small, although, it is large enough to make a noticeable bias when applying the estimation model directly to
satellite observed radiances (Ellingson et al., 1992). The corrected radiance were used in the regression analyses.

The downward longwave flux at the surface was calculated by the angular and frequency integrations of the corresponding intensities (see Eq. (2.4)). A four-point Gaussian quadrature was used to perform the integration over the angle.

Atmospheric Profiles and Radiation Data

For the development of the DLR estimation technique, we use a set of 1600 soundings compiled by Phillips et al. (1988) as input to the radiative transfer model. Each sounding includes temperature values at 65 different pressure levels from 0.1 to 1000 mb and the mixing ratios of $H_2O$ and $O_3$ in the corresponding 64 layers. The soundings were compiled from radiosondes ascents from land and ocean stations between 30˚S and 60˚N in latitude, and the soundings are equally divided between tropical (30˚S–30˚N) and mid-latitude, winter and summer conditions. The $O_3$ data were chosen by Phillips to be climatologically consistent with the temperature profiles, and the stratospheric $H_2O$ mixing ratio is assumed to be 3 ppm. The earth’s skin temperature and the surface air temperature are assumed equal.

Figs. 3.1a and 3.1b show the mean and standard deviation of the temperature profiles from the Phillips soundings for each category, respectively. The mean and standard deviation of the Phillips water vapor profiles for each category are shown in Figs. 3.2a, b, respectively. The McClatchey soundings (McClatchey et al., 1972) which are considered the standard atmospheres are plotted in Fig. 3.3 for reference. Since most of the
DLR originates from the lowest few kilometers of the atmosphere, we are especially interested in knowing the differences between the Phillips soundings and the McClatchey soundings in the troposphere.

It was found that the agreement between the mean Phillips temperature profiles in tropics and the mid-latitude winter compared with those of McClatchey is fairly good in the troposphere. However, there are consistent differences between the mid-latitude summer profiles where the Phillips mean temperatures are about 3°C lower than those of McClatchey.

For the tropospheric water vapor profiles, it can be seen that the differences are quite large. The mean Phillips water vapor mixing ratios from both seasons of tropics and from mid-latitude summer are all lower than those of McClatchey with the largest difference at the surface of about 3 g/kg. The mean water vapor mixing ratios from Phillips mid-latitude winter soundings, on the other hand, are always higher with the largest difference also at the surface of about 1 g/kg.

One might think that an estimation equation constructed from the Phillips soundings which have differences shown above will result in overestimation or underestimation of the DLR relative to the McClatchey. However, since both the upwelling clear-sky radiances and the DLR are calculated from the same soundings, they are consistent with each other. It is the variances of the soundings that are important to the ability of prediction but not the means in this case. However, there is no objective references to determine whether the variations of the Phillips soundings are representative.

For model verification, we used the analyses from output of the Global Data Assimilation System at National Meteorology Center (NMC)
which are used in the Medium Range Forecast model. Horizontally, they are distributed over 162x82 grid points with the longitudinal resolution of 2.22° and the latitudinal resolution of 2.18°. Vertically, there are 19 sigma levels. Earth’s skin temperature and surface pressure are reported for each grid as well.

The upwelling clear-sky radiances and the clear-sky DLR were calculated for four primary synoptic times for Dec. 22, 1990 by the radiative transfer model described above. The temperatures are interpolated from the 18 layer temperatures to the model pressure levels which are calculated based on the sigma levels and the surface pressure. There are some regions which have strong discontinuities at the earth’s surface when comparing the air temperatures extrapolated from the lowest two layers with the reported skin temperatures (see Fig. 3.4). Some of them may be related to the diurnal cycle (e.g., the inversion caused by the night time surface cooling over the Sahara and Saudi Arabia deserts), and the super-heated surface over the Australia desert in the local noon time. But it cannot be determined in most cases whether these large vertical temperature gradients are due to natural variations or from incorrect analyses. This temperature difference field gives a hint that there may be some errors in those regions where strong discontinuities are seen.

The skin temperatures were used in the radiation calculation for the surface. Any errors in the skin temperatures will result in the incorrect simulations of the upwelling clear-sky radiances while this error has a smaller effect on the calculations of the DLR. The DLR estimation may have errors correlated with the skin temperature errors. If those large temperature gradients are natural, we will still expect to see problems with
it because there are relatively few Phillips soundings that have the vertical temperature gradients near surface in the comparable sizes. Therefore, the sounding sample we used is not representative for these special cases. Other algorithms and information may be needed to handle the cases which have the strong near surface temperature gradients.
Fig 3.1. Mean and standard deviation of the temperature profiles from the Phillips soundings. The terms TW, MW, TS and MS refer to tropical winter, mid-latitude winter, tropical summer and mid-latitude summer, respectively.
Fig 3.2. Mean and standard deviation of the water vapor profiles from the Phillips soundings. Other symbols as in Fig. 3.1.
Fig 3.3. The McClatchey tropical (trp), mid-latitude summer (MS), and mid-latitude winter (MW) temperature (a) and water vapor (b) profiles.
CHAPTER IV
Model Development and Regression Analyses

The construction of the models is basically statistical but it is guided with the physical insights developed in Chapter II. The possible predictees in the regression models are the DLR, the emissivity or the transmissivity. The candidate predictors are the HIRS/2 radiances, or the brightness temperatures. For every model, a set of predictors was first determined. The determination of what variables should be in the model is partly physical and partly empirical. In many cases, the residual analyses suggested the addition or removal of some variables.

Stepwise regression was used to select the most important variables and to reduce the model size. The order of insertion of variables in stepwise regression is determined by using the partial correlation coefficient as a measure of the importance of variables not yet in the equation. It is equivalent to use the partial F-value instead of the partial correlation coefficient. The partial F-value is defined as the mean square (which equals to the sum of the squares since it has one degree of freedom) of the model as though the particular variable were the last variable to enter the regression equation. At every step, the partial F-values are calculated for the predictor variables not yet in the equation and the variable with the largest F value is selected for addition or rejection by comparing its F value to a preselected significance level. This is referred as a sequential F-test or ‘F to enter’.
Afterward, the overall regression is examined for significance. Meanwhile, the partial F-values are calculated for all variables currently in the equation. The variable with the lowest F value is determined to be retained in the equation according to whether the partial F-test is significant or not significant to another preselected significance level. This is referred as the partial F-test or ‘F to remove’. When no variables in the current equation can be removed and no variables can pass the sequential F-test, the process stops (see Draper and Smith, 1981). In the following regression analyses, the significance level for F to enter was chosen to be the F value at a confidence level of 0.95, i.e., $F(1, \infty, 0.95) = 3.84$. The significance level for F to remove was chosen at the same confidence level but the F value was set to 3.83 for providing some protection for predictors already admitted to the equation.

In our regressions, the final model sizes were usually not determined by the final steps of the stepwise regressions. Instead, the sizes were determined more or less subjectively. The increase of the fraction of the total variance explained (or the reduction of the residual RMS error) by the model with the increasing number of variables entered in the regression equation is usually asymptotic to a certain value. Fig. 4.1a shows a typical curve of the percentage of the total variance explained (i.e., the $R^2$) with the increasing number of variables in the model entered in sequence determined by the stepwise procedure. (Also see Fig. 4.1b for the residual RMS error.) It can be seen that in this particular plot that the explained variance does not increase significantly after the number of variables exceeds six. Although the partial F-test allowed more variables to enter the model, we subjectively chose the size of the model at six variables. Part of
the reason was that, at a certain size of the model, the estimation error attributed to the observational noise will exceed the regressional error and the addition of more variables will eventually result in a less accurate estimation (see Ellingson et al., 1989). The sizes of the reduced models were determined with the consideration of both the regressional errors and the possible noise in the predictors (see Chapter V). Many regression models have been experimented during the course of the study, the most meaningful models are listed and discussed in the following sections.

**Clear Sky Condition**

**Linear Models**

Given the DLR as the predictee and the upwelling, clear-sky radiances as the predictors, the general linear model for clear-sky conditions is written as

\[ F_o = a_o + \sum a_i f_i(N) \]  

(4.1)

where the \( f_i \)'s are functions of radiances, and they are either linear or non-linear in radiances (e.g., cross-products of radiances).

Regression analyses were performed on (4.1). For clear-sky DLR, the first linear model was composed with the radiances as the predictors, i.e.,

\[ F_o = a_o + \sum a_i N_i \]  

(4.2)

With the guide of the stepwise regression, it was determined to include five variables in the model which are the radiances from channel 5, 6, 7, 8 and 10. This model is named Model I-1, and it has a residual RMS error of 12.29 Wm\(^{-2}\) and an \( R^2 \) of 96.9%. Table 4.1 lists the regression coefficients and their standard errors. The ratio of the regression coefficient to its
standard error forms a statistic that is used to test whether the coefficient differs significantly from zero. For a significance level of 0.05, whenever the ratio is less than 1.96 (which is $t(\infty, 0.975)$), we decided to exclude that variable because its coefficient is not significantly different from zero.

The sequence of the variable selection from stepwise regression is shown in Fig. 4.1. Although it is often difficult to explain the physical reasons for the choice of variables in a statistical model, it appears that the choice of variables of this model has a physical explanation. It is known that most of the DLR come from the lowest layer of the atmosphere and is determined by the corresponding temperature and the water vapor content. The stepwise regression did choose the variables which strongly correlate with the near surface temperature or the water vapor content. The first variable chosen was the radiance from channel 8 which is centered at about 11.1 $\mu$m. It is very sensitive to the surface temperature since it is located in the 8 to 12 $\mu$m atmospheric window. The second variable chosen was the

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Regression Coeffs.</th>
<th>Std. Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-117.138</td>
<td>12.24</td>
</tr>
<tr>
<td>N5</td>
<td>-24542.0</td>
<td>611.8</td>
</tr>
<tr>
<td>N6</td>
<td>49569.4</td>
<td>1023.</td>
</tr>
<tr>
<td>N7</td>
<td>-26268.3</td>
<td>710.7</td>
</tr>
<tr>
<td>N8</td>
<td>12598.0</td>
<td>165.1</td>
</tr>
<tr>
<td>N10</td>
<td>-13389.0</td>
<td>535.0</td>
</tr>
</tbody>
</table>
radiance from the channel 10 which is centered at about 8.2 µm. It is located in the wing of the 6.3 µm water vapor band and therefore is sensitive to the lower tropospheric water vapor amount as well as the near surface temperature. Although these two variables alone have already explained the total variance of about 89%, the addition of the next four variables still made a not insignificant improvement which, with six variables, explains about 97% of the total variance. Channels 1 through 7 provide the information of the variations about the atmospheric temperature profile (see Table 2.1). This information supplements the ability of the regression equation to explain more variance.

As discussed in the Chapter II, the different portions of the DLR spectrum have different physical relations to the atmospheric parameters. It is likely that the variables needed to estimate the contribution from the different portions of the spectrum are weighted differently; or even different sets of variables are selected to interpret them. The regressions on the partitioned DLR were performed separately to examine this possibility.

The regression equation for the partition A (or the “near-opaque” spectral intervals) was determined as

$$F_{o,OP} = a_0 + a_1N_8 + a_2N_{10} + a_3N_6 + a_4N_2 + a_5N_7$$

(4.3)

where $F_{o,OP}$ is the portion of clear-sky DLR from partition A and the $a_i$’s are the corresponding regression coefficients. The variables are listed by the sequence found in stepwise regression.

The regression equation for the partition B (or the “near-transparent” spectral intervals) was determined as

$$F_{o,TR} = b_0 + b_1N_8 + b_2N_{10} + b_3N_6 + b_4N_2 + b_5N_5$$

(4.4)
where $F_{o, TR}^↓$ is the portion of clear-sky DLR from partition B and the $b_i$'s are the corresponding regression coefficients.

The regression equation (4.3) has an $R^2$ of 99.5% and a RMS error of 1.50 Wm$^{-2}$ (named Model I-2a). The regression equation (4.4) (named Model I-2b) has an $R^2$ of 94.1% and a RMS error of 11.76 Wm$^{-2}$. The Model I-2 for estimating the total DLR was defined as the sum of two separately regressed linear Models I-2a and I-2b, i.e.,

$$F_{o}^↓ = F_{o, OP}^↓ + F_{o, TR}^↓$$  \hspace{1cm} (4.5)

Model I-2 has a residual standard error of 12.61 Wm$^{-2}$ which is almost the same as that of Model I-1. Both models have fairly well distributed residual plots against the estimated DLR (see Figs. 4.2a and 4.3a). The estimated DLR from these two models deviate in the mean by 0.08 Wm$^{-2}$ and the differences scatter only by a standard deviation of 0.68 Wm$^{-2}$. This difference is so small that we conclude that the regression on partitioned DLR did not significantly improve the size or the distribution of the regressional error.

In (4.3), the first variable chosen is the radiance of channel 8 ($N_8$), and it alone explains 98.1% of the total variance and gives a residual standard error of less than 3 Wm$^{-2}$. This means that the unexplained variance and residual error in the total DLR comes mainly from the “near-transparent” spectral intervals (the partition B). The effort of adding variables through stepwise process is attributed mostly to explaining the variance from partition B where water vapor variation plays the major role. Since the DLR in partition A is commonly surface temperature dependent while the DLR in partition B is mainly modulated by the amount of water vapor, we want
to see if the regression Models I-1 and I-2 have drawn the same conclusion. Notice that since the stepwise variable-choosing sequences are nearly the same for regression Models I-1, I-2a and I-2b, we suspect that the temperature effect is overwhelming even in the regression for partition B where water vapor effect was assumed to be more influential. It is likely that the channels bearing the water vapor information are also strongly influenced by the temperature variations. Their radiance variations corresponding to the variation of the water vapor content are overshadowed by the variation of the temperature. Therefore the temperature variation dominates the model statistically and the water vapor variation is ignored. For this reason, we need to check whether the models are equally responsive to both temperature and water vapor variations.

As a measure of the model’s responsiveness, Figs. 4.2 and 4.3 show the residual as functions of the surface temperature ($T_s$) and the column precipitable water ($PW$). For both models, it can be seen that although the spread of the residual is nearly independent of $T_s$, there is a tendency of rising residual with increasing amount of precipitable water. These plots imply that both models will perform satisfactorily for an area with low precipitable water. However, for large precipitable water cases, we would expect an increasing underestimation with the increasing PW (e.g., the tropical ocean regions).

Another common feature is the large negative residual patterns ($<-25 \text{ Wm}^2$) corresponding to the high surface temperatures (see the lower right corner of plots (a) or (b) in Figs. 4.2 and 4.3). These data are from hot but relatively dry soundings. Thus, it can be expected that these two
models will systematically overestimate the DLR in desert regions where the surface temperature is high but the air is relatively dry.

The regressions on the partitioned DLR do not prevent the temperature effect from dominating the channels’ weighting. The water vapor modulation effect is not properly seen in the Model I-2b. This suggests that either non-linear functions and/or other variables are needed to take into account the water vapor variation correctly.

Assuming the problem seen above is due to the neglect of the correlation of the temperature and water vapor, we attempted a higher order model in which the radiances, their cross-products and squares were used as candidate predictors. To minimize the size of the full model, channels 1, 3, 4 and 9 were eliminated from the regression based on previous results. The Model I-3 determined by stepwise regression is composed of six variables and has an $R^2$ of 96.9% and a RMS error of 11.93 Wm$^{-2}$.

$$F_o = a_0 + a_1 N_5 N_7 + a_2 N_5 N_{10} + a_3 N_6 N_8 + a_4 N_6 N_{10} + a_5 N_7 N_8 + a_6 N_7 N_{10}$$

(4.6)

The residual plots shown in Fig. 4.4 for Model I-3 (Eq. (4.6)) is very similar to that of model I-1 with about the same size of residual standard errors (see Figs. 4.2a and 4.4a). Again, unfortunately, the residual plot against precipitable water still shows the dependency as seen in previous two models (see Figs. 4.2c, 4.3c and 4.4c). Thus, the model with non-linear predictors does not respond properly to the water vapor variation, either.

Two models based on transformed variables were also studied. Model I-4 includes the radiances and their logarithms, whereas Model I-5
uses brightness temperatures instead of radiances as predictors. Model I-4 (see Eq. (4.7)) has an $R^2$ of 96.9% and a RMS error of 12.06 Wm$^{-2}$.

$$F_o^{-\downarrow} = a_0 + a_1 N_5 + a_2 N_6 + a_3 N_7 + a_4 N_8 + a_5 \ln(N_{10}) + a_6 \ln(N_{11})$$

Model I-5 (see Eq. (4.8)) has an $R^2$ of 96.8% and a RMS error of 12.13 Wm$^{-2}$.

$$F_o^{-\downarrow} = a_0 + a_1 T_5 + a_2 T_6 + a_3 T_7 + a_4 T_8 + a_5 T_{11}$$

However, neither of them surpasses the previous models. Both models preserve the unwanted residual dependency on precipitable water in moist conditions (see Figs. 4.5 and 4.6). It suggests that the linear models, probably no matter of the degree of order or the transformations, cannot take into account the water vapor effect correctly due to the overwhelming temperature dependence in the DLR. It may be an intrinsic limitation of the linear models.

**Emissivity-approach Models**

It was shown in Chapter II that the monochromatic downward longwave flux can be expressed as a product of a blackbody emission at a certain effective temperature ($T_e$) and an atmospheric emissivity (see Eq. (2.11)). The effective temperature for the thermal emission is determined by (2.12). For interpretative purpose, (2.11) is rewritten as

$$DLR = \bar{B} \cdot \mathcal{E}$$

where the emissivity is a function of the concentrations and distributions of the atmospheric gases. Since water vapor is the most variable parameter and has the most profound effect on the DLR among the atmospheric gases, we assume that the variability of the emissivity is a function of water vapor only.
When (4.9) is referred to the total spectrum of DLR, the thermal emission $\bar{B}$ is simply equal to $\sigma T_e^4$, where $\sigma$ is the Stefan-Boltzmann constant. However, when (4.9) is applied to a portion of the DLR spectrum, then the thermal emission can be expressed as a fraction of the total thermal emission. That is,

$$\bar{B}(T_e) = g(T_e; \nu_1, \nu_2) \cdot \sigma T_e^4$$

(4.10)

where the fraction function $g$ is defined in terms of the Planck function as

$$g(T; \nu_1, \nu_2) = \frac{\int_{\nu_1}^{\nu_2} B_\nu(T) d\nu}{\int_0^{\infty} B_\nu(T) d\nu}$$

(4.11)

Analytical expression for $g$ in terms of the temperature $T$ cannot be found. The approach followed here is to perform integrations of (4.11) for several different temperatures and find the coefficients for a second-order polynomial with the least square fit.

The DLR spectrum was divided into the seven intervals as listed in Table 4.2. Four of them belong to Partition A (the “near-opaque” intervals) and the rest three belong to Partition B (the “near-transparent” intervals). In each interval, $g$ is written as

$$g(T; \delta \nu_i) = a_0 + a_1(\delta \nu_i) T + a_2(\delta \nu_i) T^2$$

(4.12)

It should be noted that although the interval 2240-3000 cm$^{-1}$ is considered as a “near-opaque” interval, it should actually cover the range 2240-2400 cm$^{-1}$ because the remaining partition is very transparent. Since the energy is so small in the interval 2400-3000 cm$^{-1}$, we simply combine it to the previous interval. Our results show that this has no significant effect on the regression of the partitioned DLR.
When dealing with only two partitions, polynomials of order three were applied. Both least square fits have an $R^2 = 100.00\%$ and the corresponding coefficients are given in Table 4.3.

<table>
<thead>
<tr>
<th>$\delta \nu_i (cm^{-1})$</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-400</td>
<td>1.18359</td>
<td>-5.60968e-3</td>
<td>7.36505e-6</td>
<td>0.9998</td>
</tr>
<tr>
<td>400-600</td>
<td>0.491011</td>
<td>-1.03046e-3</td>
<td>2.05531e-7</td>
<td>0.9997</td>
</tr>
<tr>
<td>600-700</td>
<td>0.016549</td>
<td>8.38390e-4</td>
<td>-1.82692e-6</td>
<td>0.9949</td>
</tr>
<tr>
<td>700-1320</td>
<td>-0.754333</td>
<td>6.92881e-3</td>
<td>-1.00827e-5</td>
<td>1.0000</td>
</tr>
<tr>
<td>1320-1800</td>
<td>-0.0608758</td>
<td>4.01941e-5</td>
<td>1.53759e-6</td>
<td>0.9997</td>
</tr>
<tr>
<td>1800-2240</td>
<td>0.0678630</td>
<td>-6.79731e-4</td>
<td>1.73592e-6</td>
<td>1.0000</td>
</tr>
<tr>
<td>2240-3000</td>
<td>0.0492180</td>
<td>-4.30018e-4</td>
<td>9.47198e-7</td>
<td>0.9987</td>
</tr>
</tbody>
</table>

Table 4.3. The coefficients of the polynomials for Partition A and B. Units as in Table 4.2.

<table>
<thead>
<tr>
<th>$\delta \nu_i (cm^{-1})$</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partition A</td>
<td>1.72251</td>
<td>-0.0110991</td>
<td>2.98198e-5</td>
<td>-2.64205e-8</td>
</tr>
<tr>
<td>Partition B</td>
<td>-0.708932</td>
<td>0.0109289</td>
<td>-2.90979e-5</td>
<td>2.54012e-8</td>
</tr>
</tbody>
</table>

When dealing with only two partitions, polynomials of order three were applied. Both least square fits have an $R^2 = 100.00\%$ and the corresponding coefficients are given in Table 4.3.

There is a problem in applying (4.9) to real data since we have no a priori information as to the effective temperature unless the atmospheric emissivity is known (see Eq. (2.12)). The approach followed here is to find the effective temperature empirically by experimenting with the brightness temperatures from various channels.

To gain physical insights into the variable selection, we first parameterize the emissivity in terms of the amount of the precipitable
water. As discussed in Chapter II, the emissivity is approximately linear in the precipitable water and its square (see Eq (2.20)) for the near-transparent regions.

An atmospheric emissivity for each HIRS channel is defined based on a given channel brightness temperature (assumed to be the effective temperature) by:

\[ \varepsilon_k = \frac{DLR}{\sigma T_k^4} \]  

(4.13)

where the subscript \( k \) denotes the channel number.

Since most of the DLR originates in the lowest one kilometer of the atmosphere, we expect that the effective temperature can be represented by the brightness temperature of the channels which are sensitive to the near surface fields. Fig. 4.7 shows the emissivities corresponding to the total DLR defined by the brightness temperatures from channel 8, 10 and 13 plotted as functions of the precipitable water. One notices that there are some values of emissivity exceeding one which is the upper limit by the definition of emissivity. This is because the brightness temperature used is not the actual effective temperature but an approximation. When the atmosphere becomes more optically thick (i.e., in general, more precipitable water), the effective emitting layer lowers closer to the surface which usually has a higher temperature. On the other hand, because of saturation of the absorption, the satellite-viewed radiances come from a higher emitting layer which usually has a lower temperature. The satellite’s brightness temperature and the effective emitting temperature are then in a negative correlation. This also partially explains why the curves shown in Fig. 4.7 are not quadratic in precipitable water but more like a square root of
it. As another reason, indicated by the strong line limit, the emissivity for the entire spectrum will be a function of the square root of the absorber amount (and eventually the log at high absorber amounts). Fig. 4.8 shows the same emissivities plotted against the square root of the precipitable water. They are almost straight lines.

With the preceding analysis in mind, it was assumed that the emissivity was linear in PW and its square root. Regressions were performed on the emissivities defined by brightness temperatures from three different channels. The models are in the form of

\[ F_o \downarrow = \sigma T_i^4 \cdot (a_0 + a_1 pw + a_2 \sqrt{pw}) \]  

(4.14)

where \( T_i \) is the brightness temperature from the channel \( i \). Summary results are listed in Table 4.4.

<table>
<thead>
<tr>
<th>Models</th>
<th>( T_i )</th>
<th>Flux errors (( Wm^{-2} ))</th>
<th>Emissivity regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model II-1</td>
<td>( T_8 )</td>
<td>5.31 -0.19 0.0143 96.9</td>
<td></td>
</tr>
<tr>
<td>Model II-2</td>
<td>( T_{10} )</td>
<td>3.96 0.08 0.0115 98.9</td>
<td></td>
</tr>
<tr>
<td>Model II-3</td>
<td>( T_{13} )</td>
<td>4.36 0.10 0.0130 98.0</td>
<td></td>
</tr>
</tbody>
</table>

Residual plots of the regressions of the emissivity, and the resulting flux errors as functions of surface temperature and the precipitable water for Models II-1, II-2 and II-3 are shown in Figs. 4.9, 4.10 and 4.11, respectively. All three models have very good residual plots for the
emissivity regressions. The flux errors are very evenly distributed over the entire range of the precipitable water which is an exciting feature. It means that this approach of DLR estimation has achieved a correct response to the water vapor variation opposed to the linear models discussed in last section. With the separation of the temperature and water vapor dependencies, the above model expression is successful in interpreting the temperature and water vapor variations equally. Furthermore, the resulting standard error in DLR is less than 5 Wm$^{-2}$ which is excellent in DLR estimation. These models are readily usable once an accurate estimation of precipitable water is provided.

The emissivity defined by the brightness temperature from channel 10 has the smallest regression error with precipitable water and also has the smallest flux error (i.e., Model II-2). We decided to use the emissivity defined by $T_{10}$ in the studies discussed below.

Using Model II-2 as a guideline, we developed expressions for the atmospheric emissivity in terms of satellite brightness temperature. The first attempt was to express it linearly in brightness temperatures, and the necessary coefficients were found. Model III-1 was found to be

$$F^\downarrow_o = \sigma T_{10}^d \cdot (a_0 + a_1 T_8 + a_2 T_{10} + a_3 T_{13}) \quad (4.15)$$

The emissivity residual RMS error is 0.0302 and the $R^2$ is 92.4%, while the resulting flux error has a mean of -0.04 Wm$^{-2}$ and a standard deviation of 11.07 Wm$^{-2}$. As shown in Fig. 4.12, the residual of the emissivity looks reasonable, but the flux residuals show a problem of overestimation of the hot-dry soundings.

Additional experiments were performed to estimate the emissivity by performing regressions on emissivities based on partitioned DLR. This
was done to see whether the emissivity in Partition B will be better estimated by this approach. Two models were found for the emissivities of the partitioned DLRs. Model **III-2a** and Model **III-2b** are described by (4.15) and (4.16), respectively.

\[
F_{o,OP}^{\downarrow} = g_{OP}(T_8) \cdot \sigma T_8^d \cdot (a_0 + a_1 T_7^d + a_2 T_8^d + a_3 T_{13}^d) 
\]

\[
F_{o,TR}^{\downarrow} = g_{TR}(T_{10}) \cdot \sigma T_{10}^d \cdot (a_0 + a_1 T_7^d + a_2 T_8^d + a_3 T_{13}^d) 
\]

where \(F_{o,OP}^{\downarrow}\) and \(F_{o,TR}^{\downarrow}\) are the clear-sky fluxes from the portions of the DLR in Partitions A and B, respectively, and \(g_{OP}\) and \(g_{TR}\) are the corresponding fractions of a blackbody emission.

Figs. 4.13 and 4.14 show the residual plots corresponding to Models III-2a and III-2b. In spite of the fact that the DLR in partition A is well estimated by Model III-2a, Model III-2b still has the same problem as seen in Model III-1. Apparently, a linear combination of the brightness temperatures is not good enough to represent the precipitable water in varying combination conditions, particularly for the hot-dry combination.

As an approximation for optically thin conditions, the upwelling clear-sky radiance can be expressed as the surface emission attenuated exponentially by the optical depth which is approximately proportional to the absorber amount, i.e.,

\[
N = B^* \cdot \text{Exp}(-\tau) 
\]

where \(B^*\) is the surface emission and \(\tau\) is the optical depth of the atmosphere. For two channels, where one is observed in a frequency without absorption and the other is observed in a frequency where is optically thin, the ratio of the radiances of such two channels is approximately proportional to the optical depth.
Following this principle, experiments were attempted to find the relationship of precipitable water with the HIRS radiances or brightness temperatures. It was found that the temperature ratios have better estimations of the precipitable water than that of the radiance ratios. Fig. 4.15 shows the scatter plots of the precipitable water as functions of the temperature ratios in which the ratio of temperatures from channel 13 and 10 indicates a strong correlation with the precipitable water. Adding the temperature ratios and their square roots (as representations of precipitable water and its square root) to the full model of emissivity estimation equation, Model IV-1 was determined as

\[
F_o^\downarrow = \sigma T_{10}^d \cdot \left( a_0 + a_1 \sqrt{\frac{T_{13}}{T_{10}}} + a_2 \frac{T_{13}}{T_{10}} + a_3 \sqrt{\frac{T_{13}}{T_8}} + a_4 \sqrt{\frac{T_{13}}{T_7}} + a_5 \sqrt{T_8} + a_6 T_7 + a_7 T_8 \right) \tag{4.19}
\]

which has a emissivity RMS error of 0.0245 and an \( R^2 \) of 93.7% while the resulting flux error has a mean of 0.00 Wm^{-2} and a standard deviation of 8.69 Wm^{-2}.

To further ensure that Model IV-1 responds to the water vapor variation correctly, the regression was performed for the emissivity defined with the DLR from Partition B. Model IV-2 is defined as

\[
F_{o,TR}^\downarrow = g_{TR} (T_{10}^d) \cdot \sigma T_{10}^d \cdot \left( a_0 + a_1 \sqrt{\frac{T_{13}}{T_{10}}} + a_2 \frac{T_{13}}{T_{10}} + a_3 \sqrt{\frac{T_{13}}{T_8}} + a_4 \sqrt{\frac{T_{13}}{T_7}} + a_5 \sqrt{T_8} + a_6 T_7 + a_7 T_8 \right) \tag{4.20}
\]

Figs. 4.16 and 4.17 show the residual plots for Models IV-1 and IV-2. It can be seen that the residual plots of emissivity are very evenly distributed and the residual of flux has no obvious dependency on either
surface temperature or the precipitable water. The problem related to the hot-dry conditions disappeared as well. Also it is worth mentioning that the standard error of the estimated DLR of Model IV-1 is about 9 Wm\(^{-2}\). This is the first model to estimate the clear sky DLR using information only from HIRS observations that has a standard error less than 10 Wm\(^{-2}\). Note that an error of 10 Wm\(^{-2}\) in the net monthly-average surface energy budget will lead to about a 1°C error in the ocean temperature on an annual average.

**Cloudy Sky Condition**

**Linear Models**

The general expression used in the regression analyses for cloudy-sky condition is:

\[
F_c(z_{cb}) = b_o(z_{cb}) + \sum b_i(z_{cb})f_i(N) \tag{4.21}
\]

where the coefficients \(b\)'s are functions of cloud base height.

Since there are no significant differences among the linear models for clear-sky DLR, the initial regressions for the cloudy-sky DLR were performed on the basis of Model I-1. The estimation models for cloudy-sky DLR were constructed for some certain cloud base heights, namely 240, 320, 400, 500, 600, 675, 750, 850, 900 and 950 mb. The cloudy-sky DLR from other cloud base heights will be obtained by interpolation.

A \(t\)-test for the null hypothesis \(H_0: \beta_i = 0\) against the alternative \(H_1: \beta_i \neq 0\) was also performed to eliminate the predictors which have large errors in their regression coefficients. The regressions for cloudy-sky DLR when cloud base heights are between 240 mb to 500 mb show that the \(t\)-test fails to reject the null hypothesis for the constant term. The regression
equations for cloudy-sky DLR when cloud base heights are between 600 mb to 950 mb are composed of \( \{N_5, N_6, N_7, N_8, N_{10}, N_{11}\} \) with a constant term except that \( N_{10} \) was removed from the regression equation when cloud base is at 675 mb also due to the \( t \)-test. The linear models for estimating the cloudy-sky DLR from different cloud base heights are listed in Table 4.5.

Fig. 4.18 shows the standard error and \( R^2 \) of DLR regressions for clear sky (Model I-1) and cloudy skies (Models V-1 to V-4 and VI-1 to VI-6). An interesting feature is that all cloudy sky DLR estimations have smaller regression errors than those for clear-sky condition. There are two possible reasons that caused the above results. The presence of a black cloud acts like a blanket that prevent the surface from receiving any radiation above the cloud. Thus it prevents the atmospheric variation above the cloud base from affecting the DLR. Secondly, in the less opaque spectral intervals (i.e., the partition B), the emission by the cloud, which is determined by the cloud base temperature, dominates the emission from the water vapor if the cloud is low or the atmosphere is dry (in both cases there is less water vapor to contribute to the emission and absorption of the radiation).

Fig. 4.19b shows the vertical distribution of the standard deviation of the cloud base temperature. As expected, for low clouds, the variation of the DLR in partition B follows the variation of the cloud base temperature (see Fig. 4.20b). When the cloud base rises, the DLR variation is determined by variations of both cloud base temperature and water vapor amount. An obvious feature in Fig. 4.20 is that the DLR in partition A is not affected by the presence of the clouds. It is understandable since all the DLR in partition A originates from a very thin layer above the surface due to its
opacity. It is so opaque that any changes above that surface layer have little effect on the DLR reaching the surface.

The vertical distribution of the total DLR variation is shown in Fig. 4.21, and a minimum of the standard deviation of the DLR was found at about 700 mb. The shape of the curve actually matches quite well with that of the regression error. The different sizes of the regression errors are simply due to the different degrees of variation in DLR itself.

Table 4.5. Description and regression results of Models V-1 to V-4 and VI-1 to VI-6 which estimate the cloudy-sky DLR from different cloud base heights ($P_{cb}$).

<table>
<thead>
<tr>
<th>Models</th>
<th>$P_{cb}$ (mb)</th>
<th>Parameters</th>
<th>RMS error (Wm$^{-2}$)</th>
<th>$R^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model V-1</td>
<td>240</td>
<td>N5, N6, N7, N8, N10 (*)</td>
<td>10.36</td>
<td>99.9</td>
</tr>
<tr>
<td>Model V-2</td>
<td>320</td>
<td>N5, N6, N7, N8, N10 (*)</td>
<td>8.75</td>
<td>99.9</td>
</tr>
<tr>
<td>Model V-3</td>
<td>400</td>
<td>N5, N6, N7, N8, N10 (*)</td>
<td>7.40</td>
<td>100.0</td>
</tr>
<tr>
<td>Model V-4</td>
<td>500</td>
<td>N5, N6, N7, N8, N10 (*)</td>
<td>6.26</td>
<td>100.0</td>
</tr>
<tr>
<td>Model VI-1</td>
<td>600</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>5.44</td>
<td>99.2</td>
</tr>
<tr>
<td>Model VI-2</td>
<td>675</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>5.27</td>
<td>99.2</td>
</tr>
<tr>
<td>Model VI-3</td>
<td>750</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>5.59</td>
<td>99.1</td>
</tr>
<tr>
<td>Model VI-4</td>
<td>850</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>6.74</td>
<td>98.7</td>
</tr>
<tr>
<td>Model VI-5</td>
<td>900</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>7.43</td>
<td>98.5</td>
</tr>
<tr>
<td>Model VI-6</td>
<td>950</td>
<td>N5, N6, N7, N8, N10, N11</td>
<td>7.46</td>
<td>98.6</td>
</tr>
</tbody>
</table>

where (*) denotes the linear combination without a constant term.
The characteristics of the DLR in partition A give a hint for estimating the cloudy-sky DLR by estimating the difference between the cloudy-sky DLR and the clear-sky DLR. This is approximately the same as estimating the cloudy-sky DLR in partition B (see Fig. 4.20). Since partition B is nearly optically transparent, the following transmissivity approach was performed to account for the correct modeling of the water vapor variation which was not properly modeled in the linear models (see Fig. 4.22c).

Transmissivity-approach Models

The difference between cloudy sky DLR and clear sky DLR may be written as

$$
\Delta F_v^{1}(z_c) = \pi B_v(z_c) T_{Fv}(0, z_c) - \pi B_v^{(z_c, z_t)} E_{Fv}(z_c, z_t)
$$

(4.22)

where

$$
\pi B_v^{(z_c, z_t)} \equiv \left[ -\int_{z_c}^{z_t} \pi B_v(z') \frac{\partial T_{Fv}(0, z')}{\partial z'} dz' \right] E_{Fv}(z_c, z_t)
$$

(4.23)

Since little DLR comes from the atmosphere above the cloud, the second term in (4.22) is negligible compared with the first term. We write, approximately,

$$
\Delta F_v^{1}(z_c) \approx \pi B_v(z_c) T_{Fv}(0, z_c)
$$

(4.24)

Although the transmissivity in (4.24) is only for the layer between the surface and the cloud base, we expect it to be some function of the precipitable water since most of the precipitable water is located in the lowest part of the atmosphere.
If the cloud base temperature is known, the atmospheric transmissivity may be defined as

$$T_F(0, z_c) = \frac{\Delta F^i(z_c)}{\pi B^i(z_c)}$$  \hspace{1cm} (4.25)

Fig. 4.23 shows the transmissivity defined in (4.25) for the entire spectrum as a function of the precipitable water with the cloud base at 675 mb. The scattered points in Fig. 4.23 can be well fitted by an exponential decay with precipitable water except for low precipitable water conditions. It indicates that the approximation made in (4.24) is legitimate.

Without data for the cloud base temperature, one must choose a temperature that can best represent the actual cloud base temperature (see Eq. (4.25)). Our analysis of the calculation shows that the brightness temperature of channel 10, although not always representative of the cloud base temperature, gives a good estimate of the transmissivity. The transmissivity as defined in (4.25) but for the full spectrum using the brightness temperature of channel 10 is written as

$$T_F(0, z_c) = \frac{F^i_c(z_c) - F^i_o}{\sigma T_{10}^4}$$  \hspace{1cm} (4.26)

Following the form developed for the clear-sky DLR, the regression analyses for the transmissivity (defined in (4.26)) was performed using seven different temperature terms. For example, Model VII-1 for estimating the DLR difference when a cloud base is at 675 mb is:

$$\Delta F^{i,c}_{T_{10},675} = \sigma T_{10}^4 \left( a_0 + a_1 \sqrt{T_{10}^{13}} + a_2 \frac{T_{10}^{13}}{T_{10}} + a_3 \sqrt{T_{10}^{13}} T_8 ight. \\
+ a_4 \sqrt{T_{10}^{13}} T_7 + a_5 \sqrt{T_8} + a_6 T_7 + a_7 T_8 \bigg)$$  \hspace{1cm} (4.27)
The transmissivity regression has a RMS error of 0.0193 and an $R^2$ of 89.9%. The resulting flux error has a mean of 0.00 Wm$^{-2}$ and a standard deviation of 6.73 Wm$^{-2}$.

Fig. 4.24 shows the residual plots for Model VII-1. Although the shapes of the spread of both the residual of the transmissivity and the flux error are satisfactory, this formulation automatically gives a larger error in a cloudy scene than that if the cloudy sky DLR is estimated directly.

This can be clarified by writing the total DLR regression error under a partly cloudy sky with cloud fraction $A$ as, depending on the formulation,

\[
\delta F = \sqrt{(1-A) \delta F_o^2 + A \delta F_c^2}
\]  

(4.28)

or

\[
\delta F = \sqrt{\delta F_o^2 + A (F_c - F_o)^2}
\]  

(4.29)

where $\delta F_o$, $\delta F_c$, $\delta (F_c - F_o)$ are the regression RMS errors for the clear-sky DLR, the cloudy-sky DLR, and the DLR difference, respectively. Since the clear-sky DLR estimate has a larger regression error than any cloudy-sky DLR estimate (i.e., $\delta F_o^2 \geq \delta F_c^2$), the minimum error of the DLR estimate in (4.28) equals to $\delta F_c$ for an overcast condition. However, the minimum error for (4.29) equals to $\delta F_o$ in a complete clear sky.

Fortunately, since both Models IV-1 and VII-1 behave well with the same variables, we can combine them to estimate the cloudy sky DLR directly instead of estimating the flux difference. Thus, we define the cloudy-sky transmissivity as

\[
T_{F}(0, z_c) = \frac{F_{c,z_c}}{\sigma T_{10}^4}
\]  

(4.30)
and express the transmissivity as a linear combination of the above seven temperature terms.

The t-test on the regression coefficients eliminates some variables in the models for estimating the cloudy-sky DLR with cloud base pressures less than 600 mb. The models for estimating the cloudy-sky DLR with high cloud is unchanged as in the Model IV-1. That is, when the cloud base pressures are less than 600 mb, the models are in the form shown below:

\[
F_{c,Pcb} = \sigma T_{10}^4 \cdot (a_0 + a_1 \sqrt{T_{13} \over T_{10}} + a_2 T_{13} \over T_{10} + a_3 \sqrt{T_{13} \over T_{8}}
+ a_4 \sqrt{T_{13} \over T_{7}} + a_5 \sqrt{T_{8}} + a_6 T_7 + a_7 T_8)
\] (4.31)

For cloud base pressure equals to or greater than 600 mb, the models are in the following form:

\[
F_{c,Pcb} = \sigma T_{10}^4 \cdot (a_0 + a_1 \sqrt{T_{13} \over T_{10}} + a_2 T_{13} \over T_{10} + a_3 \sqrt{T_{13} \over T_{8}} + a_4 T_8)
\] (4.32)

The regression results for regression equations (4.31) and (4.32) for different cloud base pressures are listed in Table 4.6.

Fig. 4.25 shows the residual plots for Model VIII-6 which estimates the cloudy sky DLR when the cloud base is at 675 mb. The resulting flux residual is very evenly distributed against the surface temperature and the precipitable water. The other cloudy-sky DLR models have similar performances, except for the magnitudes of the errors. The percentages of the explained variance ($R^2$) of the models and the resulting DLR standard errors are shown in Fig. 4.26 as functions of the cloud base pressure. Overall, the shape of the DLR standard error curve is similar to that of the linear models and shows a minimum at about 700 mb as well.
In summary, for clear-sky conditions, the emissivity-approach model has a standard error of DLR estimation about 3.5 Wm\(^{-2}\) less than that of the linear model. For cloudy skies, the same error for the transmissivity-approach models decreases by 1 to 3 Wm\(^{-2}\) when compared to their corresponding linear models. Furthermore, while the linear models always underestimate the DLR in very moist conditions, the emissivity and transmissivity models have evenly distributed residuals relative to the precipitable water. This implies that these models yield unbiased estimates from dry to moist conditions.

Table 4.6. Regression results of the cloudy-sky DLR estimation Models VIII-1 to VIII-10.

<table>
<thead>
<tr>
<th>Models</th>
<th>(P_{cb}) (mb)</th>
<th>Flux errors (Wm(^{-2}))</th>
<th>Transmissivity regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>Model VIII-1</td>
<td>240</td>
<td>7.32</td>
<td>-0.00</td>
</tr>
<tr>
<td>Model VIII-2</td>
<td>320</td>
<td>6.32</td>
<td>-0.00</td>
</tr>
<tr>
<td>Model VIII-3</td>
<td>400</td>
<td>5.48</td>
<td>-0.00</td>
</tr>
<tr>
<td>Model VIII-4</td>
<td>500</td>
<td>4.67</td>
<td>-0.00</td>
</tr>
<tr>
<td>Model VIII-5</td>
<td>600</td>
<td>3.94</td>
<td>-0.00</td>
</tr>
<tr>
<td>Model VIII-6</td>
<td>675</td>
<td>3.37</td>
<td>0.00</td>
</tr>
<tr>
<td>Model VIII-7</td>
<td>750</td>
<td>3.58</td>
<td>0.01</td>
</tr>
<tr>
<td>Model VIII-8</td>
<td>850</td>
<td>5.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Model VIII-9</td>
<td>900</td>
<td>5.99</td>
<td>0.00</td>
</tr>
<tr>
<td>Model VIII-10</td>
<td>950</td>
<td>6.22</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Although the advantages of the emissivity and transmissivity-approach models are obvious, it is not certain that they are more stable than the linear models. The next chapter will discuss the sensitivities of the models to various sources of errors.

**Topographical Adjustment**

In general, the DLR reaching the earth’s surface is usually lower for an elevated area due to its lower surface temperature and the drier air. The DLR estimation has to be able to distinguish the topographical difference. Since the Phillips soundings are all reduced to 1000\textit{mb} surface, the regression models derived from the Phillips sounding are theoretically only applicable to area with surface pressure of 1000\textit{mb}. Erroneous estimation will be seen when the model is applied to other surface pressures, e.g., a mountainous area.

The approach followed here was to calculate the DLR and radiances for several surface pressure levels. Regressions were performed to obtain different equations for each elevation. A surface pressure value will be needed to interpolate the DLR estimates to its surface pressure level.

First, the radiation data for the elevated area were simulated by setting the surface pressure of the soundings to a certain value. This means that the portion of a sounding lower than that pressure level is removed while everything above it is unchanged. Since it is very possible that only a few of the Phillips soundings are from elevated stations, the method to adjust the soundings to represent the elevated area may cause errors. However, this is the best we can do right now.
The regression analyses were performed for each set of radiation data for different elevations. Because the DLR usually decreases smoothly and monotonically with the decreasing surface pressure, the DLR can be interpolated from the surface pressure levels for which the DLR estimates are available. The regressions were also performed for cloudy-sky conditions with various surface pressure levels. The results are not shown because they are similar in magnitude to those for surface pressure of 1000 mb.

When the surface pressures are not available, which is the case if satellite data is the only data source, an estimate of the surface pressure is needed. The approach for estimating the surface pressure was to derive empirical equations for estimating the surface pressure from a given elevation. Since this method gives only a constant surface pressure for all areas with the same elevation, the seasonal and latitudinal dependencies were supplemented for a better estimation. The empirical estimation equations for the surface pressure were obtained using the hypsometric formula calculated with the McClatchey soundings. Representative surface pressure fields were generated with the topography through the estimation formulae for the summer and winter. These pressure fields will be used as the surface pressure estimates.
Fig. 4.1. An example of the fractional variance explained (a) and residual regression error (b) as variables are added in the stepwise regression.
Fig. 4.2. Residual plots of Model I-1. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.3. Residual plots of Model I-2. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.4. Residual plots of Model I-3. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.5. Residual plots of Model I-4. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.6. Residual plots of Model I-5. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.7. Emissivities defined by different brightness temperatures plotted against precipitable water. (a) Channel 8; (b) Channel 10 and (c) Channel 13.
Fig. 4.8. Emissivities defined by different brightness temperatures plotted against square root of the precipitable water. (a) Channel 8; (b) Channel 10 and (c) Channel 13.
Fig. 4.9. Residual plots of Model II-1. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.10. Residual plots of Model II-2. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.11. Residual plots of Model II-3. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.12. Residual plots of Model III-1. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.13. Residual plots of Model III-2a. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.14. Residual plots of Model III-2b. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.15. (a) Precipitable water vs. the ratio of brightness temperatures of channels 13 and 10; (b) Precipitable water vs. the ratio of brightness temperatures of channels 13 and 7.
Fig. 4.16. Residual plots of Model IV-1. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.17. Residual plots of Model IV-2. (a) Residual Emissivity vs. estimated Emissivity; (b) Residual Flux vs. surface temperature; (c) Residual Flux vs. precipitable water.
Fig. 4.18. Regression error and R² for DLR from clear and cloudy skies by linear models (Model I-1, V-1 to V-4 and VI-1 and VI-6) plotted with the corresponding cloud base pressure. (a) Regression standard error; (b) R².
Fig. 4.19. Statistics of cloud base temperature from 1584 Phillips soundings. (a) Mean Tcb; (b) Standard deviation of Tcb.
Fig. 4.20. Statistics of calculated DLR in partition A and B for clear and cloudy skies. (a) Mean DLR; (b) Standard deviation of DLR.
Fig. 4.21. Statistics of calculated DLR for clear and cloudy skies. (a) Mean DLR; (b) Standard deviation of DLR.
Fig. 4.22. Residual plots of Model VI-2. (a) Residual vs. estimated DLR; (b) Residual vs. surface temperature; (c) Residual vs. precipitable water.
Fig. 4.23. Transmissivity defined by Eq. (4.25) with the cloud base at 675 mb plotted as a function of the precipitable water.
Fig. 4.24. Residual plots of Model VII-1. (a) Transmissivity residual vs estimated transmissivity; (b) Flux residual vs. surface temperature; (c) Flux residual vs. precipitable water.
Fig. 4.25. Residual plots of Model VIII-6. (a) Transmissivity residual vs estimated transmissivity; (b) Flux residual vs. surface temperature; (c) Flux residual vs. precipitable water.
Fig. 4.26. Standard error of the resulting DLR and $R^2$ for the regression models (clear sky: Model IV-1; cloudy skies: Model VIII-1~10) plotted with the corresponding cloud base pressure. (a) Standard deviation of the resulting DLR error; (b) $R^2$ of the regression models.
CHAPTER V

Error analysis and sensitivity

General Discussion

One of the common characteristics of remote sensing products is the tendency to be biased. Only with a well-designed experiment or carefully matched data sets where the ground truth is provided can the technique be tuned. Even though, the tuning is usually limited to particular locations, seasons and meteorological conditions.

Although the absolute error is hard to estimate or verify, it is usually possible to estimate the relative error. Actually, in the view point of climate change study, a good estimation of the variations of the meteorological variables is more important than their exact values.

In order to know whether the DLR estimate is reliable in reflecting the natural variation while the estimate is partly attributed to noise and other errors, we need to know how large the estimation error would be and how sensitive the DLR estimation technique is to the error components.

The study by Ellingson and Gille (1978) showed that, for tropical regions, a systematic 1°C change in the temperature alone results in an about 6 Wm\(^{-2}\) change in the DLR, and a systematic 5% change in relative humidity alone results in an about 4 Wm\(^{-2}\) change in the DLR. The combined DLR error resulted from a systematic 1°C error in temperature and 10% error in relative humidity is about 10 Wm\(^{-2}\), and the DLR error increases to about 21 Wm\(^{-2}\) when the error in the relative humidity
increases to 20%. For a systematic 2°C error in temperature and 10% error in relative humidity, the DLR error is about 14 Wm⁻².

That study also showed that a 1°C random error in temperature alone results in an about 5 Wm⁻² error in the DLR, while a 5% random error in relative humidity alone results in an about 1 Wm⁻² error in the DLR. Although the DLR error due to the random error in relative humidity seems much smaller than that due to the systematic error, it should be noted that the calculations were made for 0.3 km thick layers and the error will usually become larger for a thicker layer.

The physical methods that use radiation models to calculate the DLR using satellite retrieved soundings will have errors in about the sizes discussed above. It should be noted that the accuracy of the TOVS sounding products is about 2.5°C for temperature, and is about 30% for the layer precipitable water (NOAA, 1984). Therefore, the errors of the DLR estimation by physical methods due to the errors in the satellite retrieved soundings range from about 10 to 20 Wm⁻² on average.

**Error Analysis**

There are possible errors in the stage of technique development as well as in the stage of application. The sizes of some errors are not assessable.

Techniques based on statistical methods face the same problem. That is the representativeness of the sample. The sample we used to develop the DLR estimation technique is a set of 1600 soundings compiled by Phillips (Phillips, 1988). The soundings are taken evenly from 30° S to 60° N for winter and summer. We assume they are representative of the distribution
of meteorological conditions for areas from tropics to mid-latitude. However, there are no very cold and dry soundings since the data set does not cover any latitude higher than 60°. The DLR estimation in the polar areas, in this case, is considered as an extrapolation. A less accurate DLR estimate in polar areas is likely to happen. In addition, while ozone profiles are given by climatology, incorrect DLR response to ozone concentration may happen. But since the ozone contribution to the mean total DLR is only 0.8%, the error due to inconsistent ozone concentration is small compared to other factors discussed below.

The second error source is the radiative transfer model which is used to calculate the downward and upward longwave spectra. The radiation model used in this study has been validated with line-by-line model calculations and observations that demonstrated a flux difference to within 2% (Ellingson and Serafino, 1984) Also, the OLR (outgoing longwave radiation) estimation scheme based on this radiation model using a similar technique was shown to be successful (Ellingson et al., 1992) We believe that the relative error from the radiative transfer model calculations is small. However, if the DLR estimate is biased, it may be attributed to the radiation calculations. Extensive, additional validation of the model will be performed during the US Department of Energy Atmospheric Radiation Measurements Program (ARM, 1990).

One of the largest error components is the regression error. The regression RMS errors range from 4 to 13 Wm⁻² for varying regression models shown in Chapter II. It should be noted here that, in the regressions, the predictors are assumed under control, i.e., they are error-free. However, there are observational errors due to instrument noise
which will more likely cause a larger error bound. We will discuss this in
the next section.

The regression errors for overcast conditions are smaller than that for
clear-sky conditions. However, in a cloudy sky, there are other possible
error sources. They are the cloud amount and the cloud base height.
Usually, it is hard to evaluate the accuracy of the cloud amount estimates,
either from ground observations or satellite retrievals, especially when there
are multiple layers of clouds. This is because the cloud amount estimates
from satellite retrievals are often simplified for a single layer of cloud, and
satellite observations usually cannot resolve the manners of the cloud
overlapping. Even observations from ground stations have a wide range of
errors due to human inconsistency and errors. A conservative guess for the
error of the cloud amount estimate is 10 to 30%. Note that the error of
TOVS derived cloud amount is estimated to be 20% (NOAA, 1984).

Obtaining information about the cloud base height is one of the
major difficulties for DLR estimation by any methods. It becomes more
complicated when there is a multi-layer cloud system. There are some
operational data available for cloud information, e.g., for the base height of
the lowest layer from ground observations and the radiatively-equivalent
cloud top height from TOVS retrieval. Neither of them can totally resolve
the problem. Although the US Air Force’s RTNeph (Real Time
Nephanalysis) provides the needed cloud information for multi-layer
conditions, the cloud base height estimates are not validated and are of a
questionable accuracy (K. Campana, personal communication). When this
technique is implemented in operational DLR estimation, we plan to obtain
the cloud base height either by subtracting a fixed thickness from cloud top
height or by assigning it a value according to cloud climatological data. The
collection of effects of different manners of estimating the cloud base
height is not done within this dissertation.

Both errors from cloud amount and base height estimation will
largely affect the accuracy of the DLR estimation. As illustrated in
Chapter I, the cloud emission fills up the DLR spectrum mostly in the
atmospheric window which accounts for about 20 to 30% of the total flux if
it is a blackbody spectrum. To give some idea of the magnitude, an overcast
adds about 45 to 80 Wm\(^{-2}\) more to the clear sky DLR when cloud base
height varies from 500 to 50 mb above the surface. The amount of the
additional DLR due to the presence of cloud is a function of cloud base
height and it is weighted by the cloud amount. There will be more
discussions of this in the next section.

**Sensitivity Study**

For partial cloudy conditions, assuming an infinitesimally thin,
stratified single-layer cloud with an effective cloud amount \(A\), we can write
the total DLR as the weighted sum of the DLRs from clear and cloudy
portions of sky, i.e.,

\[
DLR = (1 - A) \ F_o^+ + A \ F_c^+ (P_c)
\]

(5.1)

where the effective fractional cloud amount, \(A\), is the product of the cloud
emissivity and the absolute fractional cloud amount, \(P_c\) is the cloud base
pressure, and \(F_o^+\) and \(F_c^+\) are the downward longwave fluxes at the surface
from clear and cloudy portions of the sky, respectively. Note that this
expression can be generalized to a multi-layer stratified cloud system.
The estimate of $F_o^{\downarrow}$ is a function of the regression coefficients, the HIRS radiances (or the brightness temperatures) and the pressure of the surface. Besides these three dependent variables, the estimate of $F_c^{\downarrow}$ is also dependent on the cloud base height. So the estimate of total DLR, according to Eqn. (5.1), is a function of HIRS radiances, effective cloud amount, cloud base pressure and the surface pressure. That is,

$$DLR = f(N_i, A, P_{cb}, P_s)$$ (5.2)

The sensitivities of DLR to the cloud amount, cloud base height and surface height are the intrinsic characteristics of the DLR and are irrelevant to the estimation methods. They are evaluated by the following expressions.

The change of DLR with respect to a unit change in the cloud amount is

$$\frac{\partial DLR}{\partial A} = F_c^{\downarrow} - F_o^{\downarrow}.$$ (5.3)

The change of DLR with respect to a unit change in the cloud base pressure, $P_{cb}$, at a unit cloud amount is

$$\frac{1}{A} \frac{\partial DLR}{\partial P_{cb}} = \frac{\partial F_c^{\downarrow}}{\partial P_{cb}}.$$ (5.4)

The change of DLR with respect to a unit change in the surface pressure, $P_s$, is

$$\frac{\partial DLR}{\partial P_s} = (1 - A) \frac{\partial F_o^{\downarrow}}{\partial P_s} + A \frac{\partial F_c^{\downarrow}}{\partial P_s}.$$ (5.5)

where $\frac{\partial F_o^{\downarrow}}{\partial P_s}$ and $\frac{\partial F_c^{\downarrow}}{\partial P_s}$ are the changes of the DLR corresponding to the change of the surface pressure in clear and cloudy conditions, respectively.
The error in total DLR resulting from the error in cloud amount is expressed as

$$\delta \text{DLR} = \delta A \cdot \left\{ \overline{(F^\downarrow_c - F^\downarrow_o)} \pm \sigma(F^\downarrow_c - F^\downarrow_o) \right\}$$

(5.6)

where $\delta A$ is the error of cloud amount, $\overline{(F^\downarrow_c - F^\downarrow_o)}$ is the mean value of the DLR difference between cloudy and clear skies and $\sigma(F^\downarrow_c - F^\downarrow_o)$ is the standard deviation of the DLR difference.

Fig. 5.1a shows the mean values of the differences between DLR from cloudy and clear skies (i.e., $\overline{(F^\downarrow_c - F^\downarrow_o)}$) based on the calculations with 1600 Phillips soundings. The values represent the sensitivity of the total DLR to the cloud amount in a 100% error (see Eq. 5.3). It shows that the degree of sensitivity decreases with the increasing cloud base height. For example, for a cloud base at 500 mb and with an error of 30% in the cloud amount, the resulting error in the DLR is about 13.5 Wm$^{-2}$ (which equals to 45 Wm$^{-2}$ times 30%). Since the sensitivity varies with cloud base, the resulting error in the DLR is about 23 Wm$^{-2}$ if the cloud base is at 900 mb. When the error in cloud amount drops to 10%, the errors in DLR become about 4.5 and 8 Wm$^{-2}$ for the cloud bases at 500 and 900 mb, respectively. Fig. 5.1a also shows the sensitivity of DLR to the cloud amount for different seasons and latitudinal zones. Among them, the mid-latitude winter has the largest sensitivity and the tropical summer has the smallest one. This is related to the amount of water vapor in the atmosphere. In general, the drier the atmosphere, the ‘cleaner’ the atmospheric window, and, therefore, the greater effect on increasing the DLR for adding a cloud.

Fig. 5.1b shows the standard deviation of the DLR differences between cloudy and clear conditions. It represents the variability of the sensitivity of the DLR to the cloud amount error. In contrast to the mean
values of the sensitivity, the tropical summer cases have the largest
variability while the mid-latitude winter cases have the smallest one. The
results in Fig. 5.1a and 5.1b show that the mean differences dominate the
total error estimate.

The error in the total DLR resulted from the error in the cloud base
height is written as

$$\delta_{DLR} = A \cdot \delta_{P_{cb}} \cdot \left\{ \left( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \right) \pm \sigma \left( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \right) \right\} \quad (5.7)$$

where $\delta_{P_{cb}}$ is the error of cloud base height in mb, \( \left( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \right) \) and 
$\sigma \left( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \right)$ are the mean and the standard deviation of 
$\frac{\partial F_c^{\perp}}{\partial P_{cb}}$, respectively.

Fig. 5.2a shows mean values of the DLR sensitivity to the error in the
cloud base height. The values shown in the plot are for overcast condition
with an error of 100 mb in cloud base height, i.e., \( \left( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \cdot 100 \right) \). Again,
the error is a function of the cloud base height. The sensitivities range from
about 7 to 10.5 Wm$^{-2}$ for all 1600 soundings. Among the season/latitude
categories, the largest variation is seen from the minimum at about 6 Wm$^{-2}$
for low cloud to the maximum at about 12 Wm$^{-2}$ for a cloud at 550 mb in the
mid-latitude winter cases. Note that the plot is made for a cloud with 100%
coverage and 100 mb error in the cloud base height, the values read out
from the plot should be scaled by the cloud amount and the base height
error in other cases.

Fig. 5.2b shows the standard deviation of the \( \frac{\partial F_c^{\perp}}{\partial P_{cb}} \) which
indicates the range of the DLR sensitivity variation to the errors in the cloud
base heights. It is also for overcast condition with 100 mb error in the cloud
base height. Noticeably, the values shown for the low clouds are rather
large and they are relatively less variable in the range of about 3 to 4 Wm\(^{-2}\) once the cloud base pressure is less than 800 mb. Another feather is that, for low cloud, the DLR sensitivity to cloud base height for mid-latitude winter is about twice as variable as that for tropical summer. An analysis of the soundings shows this is due to the inversion in the low troposphere, and there are more inversions in mid-latitude winter than in tropical summer.

Figs. 5.3a and 5.3b show the mean and standard deviation of \(\partial F_o/\partial P_s\) multiplied by 10. The value of \(\partial F_o/\partial P_s\) represents the sensitivity of the clear sky DLR to the 10 mb error in the surface pressure. The mean sensitivity is about 4 Wm\(^{-2}\) by a 10 mb error in estimating the surface pressure and with the variability of 1 to 2 Wm\(^{-2}\). It also shows that the DLR in the tropics is more sensitive to the error in the surface pressure.

Tables 5.1a and 5.1b show the mean and standard deviation of \(\partial F_c/\partial P_s\) multiplied by 10 for all soundings. The value of \(\partial F_c/\partial P_s\) represents the sensitivity of the cloudy sky DLR to the 10 mb error in the surface pressure. It can be seen that, no matter where the surface is, the higher the cloud, the larger the sensitivity. But the differences of the magnitude of the sensitivity among different cases are on average only a few 10th of 1 Wm\(^{-2}\).

Fig. 5.4 shows the mean sensitivities of the DLR to the 10 mb error in the surface pressure for both clear and cloudy conditions. The cloud base in this case is at 300 mb. Since the sensitivity for the total DLR is a cloud amount weighted sum of the clear and cloudy sensitivities, the size of the sensitivity for the total DLR lies in between the two curves, i.e., between about 2.5 and 3.5 Wm\(^{-2}\).
The sensitivities discussed above are all intrinsic to the DLR estimation problem for both physical and statistical methods. Besides them, those associated with the DLR estimation technique developed here are noise in the observations and errors from the clear-column-radiance retrievals.

In the regressions, we considered the predictors, which are the upwelling radiances, to be under control, i.e., noise free. We obtained the regression coefficients based on this assumption. In reality, the upwelling radiances observed by a satellite instrument are always contaminated by instrument noise. To take this into account, we generated several independent sets of random noise for the radiances based on the nominal instrument noise (see Table 2.1), put them back into the regression model, and then calculated the RMS errors. The increase in the RMS error with noise contaminated radiances represents the sensitivity of the regression model to the observational noise. Putting noise in one channel and keeping the rest noise free can show the sensitivity to the noise in that particular channel.

Fig. 5.5 shows the RMS errors of the regression model IV-1 with noise free condition, noise in individual channels, and in all channels for ten noise realizations. Fig. 5.6 shows the mean and standard deviation of the RMS error for regression model IV-1 from noise-free and ten noise realizations. It is easily seen that this model is much more sensitive to noise in channel 13 than it is to noise in the other channels. Overall, for model IV-1, the instrument noise contributes about an additional 0.5 Wm$^{-2}$ to the RMS error.
Fig. 5.7 shows similar information as in Fig. 5.6 but for model I-1. Among the 5 channels used in the model, channel 6 is extremely sensitive to the noise. Channel 6 alone adds about an extra of 5 Wm\(^{-2}\) to the RMS error while noise is considered. In this model, channel 8 and 10 are very insensitive to the noise and the RMS error becomes even smaller while the noise is put into channel 8. Overall, the RMS error becomes about 19 Wm\(^{-2}\) which indicates an increase of about 6 Wm\(^{-2}\) compared to the noise free condition. This is a proof that regression model IV-1 is better than model I-1 in this regard.

It should be noted that this is the error for one radiance observation. For monthly means for the 2.5°x2.5° grid boxes, the error due to instrument noise drops to about 1/30 of the values shown because there are about 900 radiance observations in a box and the noise is considered random.

Fig. 5.8 shows the RMS errors for clear and cloudy skies for radiance with and without noise. Regression model IV-1 is used for clear sky condition and models VIII-1 to VIII-10 are used for cloudy sky conditions. It can be seen that the differences of the RMS errors are within one Wm\(^{-2}\), and the differences generally increase with the decreasing cloud base pressures.

Since this DLR estimation technique relies on the clear-column radiances, it subjects to the error in the retrieval processes if clear-column radiances are not observed directly. It should be pointed out that this error affects almost all methods that derive DLR from satellite observations including physical methods which use the retrieved profiles as input. The accuracy of the retrieval processes for obtaining clear column radiances is not reported before, but is left for future research.
The combined DLR error can be evaluated by assuming all the error components are independent and determining the square root of the sum of the squares of the errors from all components, i.e.,

\[
\delta_{DLR} = \left( \sum_i \varepsilon_i^2 \right)^{1/2}
\]  

(5.8)

where the \( \varepsilon_i \)'s denotes the errors components which are the DLR errors due to the regression, uncertainties in cloud amount, cloud base height and surface pressure.

The DLR errors for different degrees of uncertainties in cloud amount and cloud base height are shown in Fig. 5.9. Keeping in mind that the regression errors with noise range from about 3.5 to 9 Wm\(^{-2}\), it can be seen that the combined DLR error will be dominated by the error due to cloud amount uncertainty when the cloud amount uncertainties are larger than about ±30%. It should be noted here that the DLR error due to error in surface pressure is less than 4 Wm\(^{-2}\) per 10 mb error. This error could become large as well if the surface pressure estimates are not accurate.

Figs. 5.10 to 5.11 show the combined DLR errors due to errors from regression (noise included), uncertainties of cloud amount and cloud base height at various cloud amounts. The results in Figs. 5.10a and 5.10b were calculated for a 10% of the cloud amount uncertainty with ±50 mb and ±100 mb of cloud base height uncertainties, respectively. Figs. 5.11a and 5.11b were calculated for a 30% of the cloud amount uncertainty also with ±50 mb and ±100 mb of cloud base height uncertainties, respectively. There are some common features in these four plots. First, the combined DLR error are always larger when the cloud amount is larger although the cloudy conditions have smaller regression errors (see Fig. 5.8). Also noticed is that the DLR errors are generally larger when the cloud bases are lower.
When the cloud amount uncertainty equals about 30%, there are relatively very little differences among the DLR errors for different cloud amounts and cloud base height uncertainties (see Fig. 5.11a,b).

Overall, the combined DLR errors, excluding the contribution of the error in the surface pressure estimation, range from about 7 to 12 Wm\(^{-2}\) when there are ±10% uncertainties in cloud amounts and ±100 mb uncertainties in cloud base heights. When the cloud amount uncertainties rise to 30%, the range of the combined DLR error increases to about 10 to 25 Wm\(^{-2}\).
Table 5.1. Sensitivities of the cloudy-sky DLR due to the error in the surface pressure. Values shown in the table are for 10 mb uncertainties in the surface pressures with varying cloud base pressures and surface pressures. Tables (a) and (b) are for the mean values and standard deviations, respectively.

(a) Mean of \(\frac{dF_c(P_s)}{dP_s} \times 10 \text{ mb (Wm}^{-2}\text{)}\)

<table>
<thead>
<tr>
<th>Pcb (mb)</th>
<th>950</th>
<th>850</th>
<th>750</th>
<th>650</th>
<th>550</th>
<th>450</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>2.79</td>
<td>2.75</td>
<td>2.65</td>
<td>2.60</td>
<td>2.48</td>
<td>2.13</td>
</tr>
<tr>
<td>400</td>
<td>2.53</td>
<td>2.48</td>
<td>2.36</td>
<td>2.27</td>
<td>2.05</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>2.29</td>
<td>2.23</td>
<td>2.08</td>
<td>1.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>2.07</td>
<td>1.90</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>1.86</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>1.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Std Deviation of \(\frac{dF_c(P_s)}{dP_s} \times 10 \text{ mb (Wm}^{-2}\text{)}\)

<table>
<thead>
<tr>
<th>Pcb (mb)</th>
<th>950</th>
<th>850</th>
<th>750</th>
<th>650</th>
<th>550</th>
<th>450</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1.55</td>
<td>1.14</td>
<td>0.85</td>
<td>0.68</td>
<td>0.56</td>
<td>0.52</td>
</tr>
<tr>
<td>400</td>
<td>1.48</td>
<td>1.07</td>
<td>0.79</td>
<td>0.61</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1.43</td>
<td>1.01</td>
<td>0.75</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>1.39</td>
<td>0.97</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>1.36</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 5.1. Sensitivities of the DLR to a 100% error in the cloud amount. The categories ALL, TW, MW, TS, MS represent results for all soundings, tropical winter, mid-latitude winter, tropical summer and mid-latitude summer, respectively. Panels (a) and (b) are for the mean values and standard deviations, respectively.
Fig. 5.1. Sensitivities of the DLR to a 100% error in the cloud amount. The categories ALL, TW, MW, TS, MS represent results for all soundings, tropical winter, mid-latitude winter, tropical summer and mid-latitude summer, respectively. Panels (a) and (b) are for the mean values and standard deviations, respectively.
Fig. 5.2. Sensitivities of the DLR due to the errors in the cloud base heights for different average atmospheric conditions (see caption in Fig. 5.1). Panels (a) and (b) are for the mean values and standard deviations with 100 mb uncertainties in the cloud base heights, respectively.
Fig. 5.3. Sensitivities of the clear-sky DLR due to the errors in the surface pressures. Panels (a) and (b) are for the mean values and standard deviations with 10 mb uncertainties in the surface pressures, respectively.
Fig. 5.4. Mean sensitivities of the DLR due to the errors in the surface pressures in clear and cloudy-sky conditions with 10 mb uncertainties in the surface pressures. Cloud base pressure is chosen at 300 mb.
Fig. 5.5. RMS errors for regression model IV-1 for clear-sky conditions with nominal instrument noise included. Results are shown for individual channels and all channels with ten independent noise realizations.
Fig. 5.6. As in Fig. 5.5, but for average and standard deviation.
Fig. 5.7. As in Fig. 5.6 but for model I-1.
Fig. 5.8. RMS errors for clear and cloudy-sky conditions (regression model IV-1 and VIII-1 to VIII-10) with and without nominal instrument noises.
Fig. 5.9. DLR errors due to errors from uncertainties in cloud amounts and cloud base heights.
Fig. 5.10. Combined DLR errors contributed from regression error, and the errors from the uncertainties in cloud amounts and cloud base heights. Curves are shown for the indicated cloud amounts. Different degrees of uncertainties in cloud parameters are considered: (a) $\delta A = \pm 10\%$ and $\delta P_{cb} = \pm 50$ mb, (b) $\delta A = \pm 10\%$ and $\delta P_{cb} = \pm 100$ mb.
Fig. 5.11. As in Fig. 5.10 but for (a) $\delta A = \pm 30\%$ and $\delta Pcb = \pm 50$ mb, (b) $\delta A = \pm 30\%$ and $\delta Pcb = \pm 100$ mb.
CHAPTER VI
Validation

Data Description

Ground-based DLR observations are very limited in time and space coverage nowadays, because few weather stations have the routine DLR observations. As a result, global DLR field based on ground observations does not exist. Several field experiments have ground observations of the DLR matched with simultaneous sounding and satellite observations, e.g., the First International Radiation Experiment (FIRE) and the Baseline Surface Radiation Network (BSRN) (personal communication with Dr. Ellingson). Most recently, the Atmospheric Radiation Measurements (ARM) program (see ARM, 1990) has planned dedicated radiation measurements with simultaneous sounding and satellite observations. These data can be used to validate the DLR estimation technique over limited ranges. However, these experiments only covers a local area and the geography and climates within these area are not too diversified.

For the purpose of validation, extended types of climates should be included. We are especially interested to see how well the statistical technique will do when applied to regions outside of the original sounding samples. For now, the best we can do is to calculate the global DLR field with a radiative transfer model and compare it to the estimated DLR field using the developed technique with either the simultaneous satellite observations or the simulated satellite observations. Based on availability, we used the simulated satellite observations to estimate the DLR.
The DLR and HIRS radiances were calculated by Ellingson’s radiative transfer model with the analysis fields from output of the Global Data Assimilation System at the National Meteorology Center. The analysis field, in a global coverage, has a 2.22° by 2.18° grid resolution (162 x 82 grid points) and 19 vertical sigma levels at four synoptic hours each day. It also includes the earth’s skin temperature and surface pressure. The field we used is at the 00Z on the 22nd of December, 1990.

Since the DLR estimation is surface pressure dependent, we have to estimate the surface pressure while it is not available. In fact, from satellite observations alone there is no way to report the surface pressure. To limit ourselves to the satellite observations, the surface pressure estimation is essential. We obtained the elevation to pressure conversion formulae based on the hypsometric formula with 5 McClatchey soundings ranging from tropics to sub-arctic for the summer and winter. Representative surface pressure fields were generated with the topography through the conversion formulae for the summer and winter and are used as the surface pressure estimates.

**Validation Results**

Figs. 6.1 and 6.2 are the contour plots of the skin temperature and total precipitable water (PW), respectively, from the NMC analysis for 00Z Dec. 22, 1990. They are the two most important fields because most features in the clear-sky DLR field may be traced back to them. Fig. 6.3 shows the radiation model calculated clear-sky DLR field which is assumed as the ‘ground truth’ for this validation study. Similar patterns can be readily recognized from Figs. 6.1 and 6.3, such as, near-zonal distribution of DLR
over the oceans, topographical features, and the cold continents in the northern hemisphere. In spite of the complexity of the precipitable water field, similarities can still be seen when compared with the DLR field. For example, the 400 Wm$^{-2}$ DLR contour over the Indian Ocean can be linked to the 4 cm contour of the precipitable water; the 400 Wm$^{-2}$ DLR contour over northern part of the South America can be linked to the 5 cm PW contour; and most structures of DLR seen in the Pacific Oceans are combined results of the surface temperature modulated with the precipitable water.

The statistics for the latitudinal zones make the connections even obvious. Figs. 6.4a,b show the zonal mean and standard deviation of the DLR, respectively. The symmetry of the DLR to the equator can also be seen in the surface temperature and precipitable water fields (see Figs. 6.5a and 6.6). The three maxima of the standard deviation of the DLR at about 30°N, 15°S and 75°S seen in Fig. 6.4b are the results of the land/sea distribution and as well as the topography over the land (see Figs. 6.5b and 6.7).

Based on the size of the regressional error and the results of sensitivity studies, we decided to use the model IV-1 for estimating the clear-sky DLR. Fig. 6.8 shows the estimated DLR field with the surface pressure from NMC analysis and Fig. 6.9 shows the estimated DLR field with the estimated surface pressure. The estimated DLR field from both Figs. 6.8 and 6.9 show a great deal of agreement to the calculated one in Fig. 6.3 in a general sense. The estimated DLR fields, to a great extent, duplicate most of the recognizable features described in Fig. 6.3, including the topography, near-zonal pattern in oceans, and so on.
The difference fields between the estimated and the calculated DLR are shown in Figs. 6.10 and 6.11 for NMC and estimated surface pressures, respectively. The contour interval is 10 Wm$^{-2}$ excluding 0 contour for visualization reason. Positive numbers indicate overestimates and negative ones mean underestimates. It can be seen that the contours exceeding ±10 Wm$^{-2}$ are relatively sparsely distributed and no obvious structures can be observed. Over 162x82 grid points, about 75% of them are within ±10 Wm$^{-2}$ and about 91% of them are within ±15 Wm$^{-2}$ (see Fig. 6.12).

The zonal mean DLR differences shown in Fig. 6.13 reveal that, for both estimates, they agree to within ±5 Wm$^{-2}$ in the tropics and mid-latitudes but the agreement becomes worse toward both polar regions. However, since the differences between the estimated and the calculated DLR in the two polar regions are of opposite signs, these large mean differences may be caused by different effects.

To further visualize the comparison, Figs. 6.14 and 6.15 show the scatter plots of the estimated versus the calculated DLR. The correlation coefficients are as high as 0.996 for both plots. The largest deviations occur where the DLR is below 100 Wm$^{-2}$. Most of them are overestimates. Beside this, the spread along the diagonal line is very homogeneous in both plots. Note that the estimation made with estimated surface pressure have fewer underestimation in the low flux range than the one using NMC surface pressure. The points that show large underestimation on the lower left corner in Fig. 6.14 were found to be the Ellesmere Island, Canada (the neighboring island to the west of Greenland). The surface pressure field (Fig. 6.16) shows a low pressure center (< 850 mb) whose location coincides with that of the large underestimation seen in the DLR difference field (see
Fig. 6.10). Although this low pressure was not identified as being due to the elevation alone or a weather system, the underestimation gives a hint that there is a potential error in the regression model or in the topographical adjustment process.

As is usual in the regression studies, we check the regression residuals against the surface temperature and the precipitable water to ensure a correct modeling of the proper influences of those parameters. Figs. 6.17a and 6.17b show the scatter plots of DLR differences against the NMC surface temperature and precipitable water, respectively. The DLR difference shows some dependence on the surface temperature. It is consistently underestimated in the range of about 250 to 280 °K, while it is mostly overestimated in the range of about 280 to 295 °K. Some very large overestimation are seen in the low surface temperatures as well in the range of about 280 to 300 °K. These overestimates are all associated with low precipitable water (see Fig. 6.17b). This may be caused by the implicit overestimation of the precipitable water in the low precipitable water conditions, and hence, the overestimation of the emissivity, by the non-linear predictors in model IV-1.

Very low water vapor content is more difficult to infer. Furthermore, the lowest precipitable water in the Phillips soundings is about 0.15 cm. About 12% of the grid points from the NMC analysis have a precipitable water smaller than 0.15 cm. Most of them have an estimated DLR which agrees with the calculated one to within ±15 Wm⁻² as in other ranges of the precipitable water. As a result, the estimation model has included the influence of the water vapor content correctly and this is considered as a major success of the estimation model.
Scatter plots for DLR differences against the surface temperature and precipitable water are also produced for estimated DLR estimates made with the estimated surface pressure (see Figs. 6.18a, b). Fig. 6.18a presents the similar problems as seen in Fig. 6.17a but nothing apparently worse. Fig. 6.18b shows the same degree of satisfaction as in Fig. 6.17b in the aspect of the precipitable water dependence.

The last test includes the topographical adjustment and the surface pressure estimation. The DLR is elevation dependent. Generally, the higher the elevation, the lower the DLR because of lower temperature and precipitable water. Without the topographical adjustment, the DLR estimates will be generally too high. Fig. 6.19 shows the DLR difference as a function of the surface pressure. Its purpose is to see whether the adjustment process has removed the elevation dependency of the DLR difference. It can be seen that the DLR difference has no large correlation with the surface pressure (the correlation coefficient of 0.24). However, it is easily seen that DLR consistently underestimated by about 10 Wm\(^{-2}\) in most elevated areas while there are some widely spread points of overestimation connected to the relatively drier profiles. Unexpectedly, the DLR difference shows a better distribution with the surface pressure when the surface pressure is estimated from the elevation (see Fig. 6.20). The correlation coefficient is 0.12. Although it is not sure what makes the differences in the DLR estimates using the NMC and estimated surface pressures, the results indicated that the elevation-derived surface pressure can be used in the DLR estimation to a reasonable error range.

Overall, the mean DLR differences are -0.87 and 0.84 Wm\(^{-2}\) for the estimations made with NMC and estimated surface pressure fields,
respectively. The corresponding standard deviations of the DLR differences are 9.15 and 9.12 Wm\(^{-2}\). When the grid values are weighted by the area of the grid boxes, the mean differences become -0.03 and 1.75 for the NMC and estimated surface pressures, respectively.

The two-tailed Student-\(t\) test on the mean of DLR differences at the 0.05 level shows that the absolute value of a mean difference larger than 0.15 will lead us not to reject the null hypothesis, \(H_0: \mu = 0\). Based on this, we say that both estimated DLR present a “bias” in the estimation although the values are small. When area weighting is considered, the estimation made with NMC surface pressure does not present a bias.

To test whether the variance of the difference of the DLR validation fields is larger than the variance of the regressiveional error, we use the one-tailed F-test with the null hypothesis \(H_0: \sigma_1 = \sigma_2\) against the alternative \(H_1: \sigma_2 > \sigma_1\), where the indices 1 and 2 represent for the regression error and the DLR difference from validation fields, respectively. For the DLR estimated with the NMC surface pressure, \(F = (s_2^2/n_2)(s_1^2/n_1)^{-1} = (9.15^2/13284)(8.69^2/1584)^{-1} = 0.13\). For the DLR estimated with the estimated surface pressure, \(F = 0.13\). Neither of these two F values is greater than the \(F(0.95, 13284, 1584) = 1\), that means we cannot reject the \(H_0\) in either case. In another word, the variances of the error found in the DLR validation fields are not significantly greater than the error seen in the regression. However, we need to have more DLR validations over different times in order to claim the applicability of this DLR estimation model in both space and time. Nevertheless, these initial tests give considerable confidences as to its applicability.
Fig. 6.4. Zonal statistics of the calculated (solid) and estimated (dashed and dotted) DLR (Wm$^{-2}$) at 00Z, 12/22/90. (a) zonal mean, (b) zonal standard deviation.
Fig. 6.5. Zonal statistics of the surface temperature (°K) from the NMC analysis at 00Z, 12/22/90. (a) zonal mean, (b) zonal standard deviation.
Fig. 6.6. Zonal mean (solid) and standard deviation (dotted) of the column precipitable water (cm) from the NMC analysis at 00Z, 12/22/90.
Fig. 6.7. Zonal mean (solid) and standard deviation (dotted) of the topography (km) where sea surface height is assumed at 0.
Fig. 6.12. Histogram for the percentage of the total grid points (162x82) falling into the given ranges of the DLR differences for 00Z, 12/22/90.
Fig. 6.13. Zonal mean estimated minus calculated DLR for 00Z, 12/22/90. The differences associated with the NMC and estimated Ps are plotted by solid and dashed curves, respectively.
Fig. 6.14. DLR estimated using the NMC surface pressure versus the calculated DLR for 00Z, 12/22/90.
Fig. 6.15. As in Fig. 6.14 but with estimated surface pressure plotted against the calculated DLR for 00Z, 12/22/90.
Fig. 6.17. (a) DLR difference versus surface temperature, and (b) DLR difference versus precipitable water for 00Z, 12/22/90. DLR is estimated with the NMC surface pressure.
Fig. 6.18. As in Fig. 6.17 but for the estimated surface pressures.
Fig. 6.19. DLR difference versus the NMC surface pressure for 00Z, 12/22/90.
Fig. 6.20. As in Fig. 6.19 but for the estimated surface pressures.
CHAPTER VII
Summary, Conclusions and
Suggestions for Future Research

The surface energy budget plays an important role in determining many atmospheric and oceanic processes, especially on the global scale. One needs a good estimation of surface energy budget to explain some large scale phenomena. The downward longwave radiation (DLR) at the earth’s surface is one of the necessary components in the study of surface energy budget. It is very difficult to study the global distribution of the DLR using ground observations because observations are sparse or non-existent, particularly over the oceans. Although earth-orbiting satellites are the best tools for obtaining global data, the global analysis of surface radiation budget based on satellite observations is not well-developed, especially for the DLR. A comprehensive review of the methods using satellite observations for studying the surface radiation budget was done by Schmetz (1989). Many efforts have been devoted to develop methods to estimate the DLR using satellite observations but there are difficulties and drawback for each of them. In this study, a new technique was developed statistically for estimating the DLR that it overcomes many of the shortcomings of previous techniques.

The data used in the model development and validation were simulated with a detailed radiative transfer model originally developed by Ellingson and Gille (1978) since observations were not available. The input
atmospheric profiles for the radiation calculations are the 1600 Phillips soundings and the analysis fields from an NMC forecast model, respectively. The Phillips soundings (Phillips et al., 1988), which are used for model development, cover the tropics and mid-latitudes geographically and they are equally divided into winter and summer. For model validation, the NMC analysis fields provide the atmospheric profiles with global coverage which is preferred for the validation purpose since the geographical and climatological variations can be examined.

There are basically two categories of regression equations developed in this study. The linear models are the linear combinations of the functions of radiances. These functions may be non-linear in radiances, e.g., cross-products. The other group of models used the emissivity or transmissivity approaches where the definable emissivity or the transmissivity are the predictees instead of the DLR. It was found that the emissivity or the transmissivity models are superior than the linear ones in the aspects of the RMS errors of the DLR estimates as well as in the residual spread patterns.

The experiments dealing with the DLR partitioned according to opacity indicated that the sizes and distributions of the errors of the regressions were not improved significantly. The results showed that, without proper transformations of the radiances, the information of the water vapor content embedded in the radiances is overshadowed by the temperature dependence. The information of the water vapor content was not revealed in the regression until the radiance ratios were employed. These transformation successfully modeled the water vapor variation in most conditions. Along with the emissivity approach, the dependencies of
the DLR on the temperature and the water vapor amount were finally modeled correctly.

The regression RMS errors of the linear models range from about 5 to 13 Wm$^{-2}$. The RMS errors of the emissivity and transmissivity models range from about 3 to 9 Wm$^{-2}$ and are about 1 to 4 Wm$^{-2}$ lower than those of the linear models for clear and cloudy conditions. In clear-sky conditions, when the effects of noise are included, the RMS error of the linear model increases by about 6 Wm$^{-2}$ while that of the emissivity model increases by only 0.5 Wm$^{-2}$. For both clear and cloudy conditions, when instrument noise is included, the regression RMS errors of the emissivity and transmissivity models range from about 4 to 9 Wm$^{-2}$.

Sensitivity studies were performed on possible error sources. The major error components are the cloud amount, cloud base height, and the surface pressure. Among them, the DLR is the most sensitive to the error in the cloud amount on average. Overall, the combined DLR errors, excluding the contribution of the error in the surface pressure estimation, range from about 7 to 12 Wm$^{-2}$ when there are ±10% uncertainties in cloud amounts and ±100 mb uncertainties in cloud base heights. When the cloud amount uncertainties rise to 30%, the combined DLR error ranges from about 10 to 25 Wm$^{-2}$.

The model for estimating clear-sky DLR was validated using simulated radiation data. Since the DLR estimation technique needs the surface pressure as a parameter for topographical adjustment, two fields of the DLR estimates were calculated based on the NMC reported surface pressure and the elevation-derived surface pressure. The estimations are compared with the calculated DLR from the radiation model. The values of
the standard deviation of the DLR differences are about 9 Wm\(^{-2}\) using either surface pressure field. The mean differences, however, are about -1 and 1 Wm\(^{-2}\) using NMC and elevation-derived surface pressures, respectively. There are some large overestimation of DLR found to be related to relatively dry air in some regions. This is due to the improper representation of the precipitable water amount by the ratios of the brightness temperatures in the regression model. In the aspect of topographical adjustment, although dependence of the flux errors on the surface pressure seems to be removed, the DLR was underestimated consistently by about 10 Wm\(^{-2}\). This problem may be caused by improper simulation of the DLR for elevated regions.

In summary, the DLR estimation models using HIRS radiance observations have been developed for both clear and cloudy conditions. Sensitivity studies show the possible error ranges in the application with real data. Although only the five most important gases were included in the radiation calculations, the approach can be modified to allow for variations in the amount and distribution of additional gases and/or aerosols. The best models were determined to be the ones using emissivity and transmissivity approaches, i.e., basic physics fitted with statistical analyses. Compared to the linear models, they have smaller regression RMS errors and are less sensitive to the observation noise. The validation of the clear-sky DLR estimation with simulated data shows satisfactory results that the errors of the DLR estimates are in a reasonable range and the standard deviation of the errors is about 9 Wm\(^{-2}\). It should be noted here that this error estimate is for the instantaneous observation. The size of the error will decrease by a magnitude proportional to the inverse of the
square root of number of observations if the scenes vary randomly. For example, there are about 25 HIRS observations (each is obtained from a 3x3-pixel array) falling into a 2.5°x2.5° grid box. With two orbits a day, there are about 900 observations in total for one 2.5°x2.5° grid box in a month assuming 60% of chance that the clear column radiance can be obtained. For the monthly averaged DLR estimate, the error will become about thirty times smaller than that of the instantaneous DLR estimate. However, since the weather over some regions are highly persistent, e.g., the sub-tropical highs, the decrease of the error estimate may not apply in those regions. Although, the error attributed to the instrument noise will always be reduced by about that magnitude because the instrument noise is generally considered random.

The validation also reveals some problems related to the large or consistent DLR deviations. One of these problems is associated with the relatively drier air particularly for high surface temperature areas. Another problem is related to the accuracy of the surface pressure and the topographical adjustment process. Also noticed is the deficiency of the DLR estimation with the presence of inversions or super-heated surface. This suggests that more studies are needed for the simulation of radiation data for elevated areas, and for the topographical adjustment process. Also needed is to find an additional mechanism to handle the inversion and the strong vertical temperature gradient near the surface.

Since the NOAA polar orbiting satellites only observe two times a day for most area (except high latitudes where the satellite swaths overlap), there are possible sampling errors for the DLR estimation due to the non-sinusoidal diurnal variation. This sampling problem is generic to all
techniques using satellite data to estimate the DLR and should also be studied in the future.

For cloudy skies, additional cloud data are needed, including the cloud base height and the effective cloud amount. The DLR estimated with different cloud data sources, however, has not been examined. It is the most important study remaining before this technique can be used operationally. Since the sensitivity study shows that the DLR error due to the uncertainty of the cloud amount usually dominates the other error components, it is believed that an accurate estimate of the cloud amount is more important for cloudy-sky DLR estimation. However, both errors attributed to the uncertainties of the cloud amount and cloud base height are not negligible compared to the total DLR estimation error.

Despite these limitations noted above, our analysis indicates that the physical statistical method developed in this study should provide estimates of the DLR to within about 10 Wm$^{-2}$ RMS directly from the HIRS observations if the cloud properties can be accurately estimated. Furthermore, it is expected that the errors will be considerably smaller for monthly average, except where persistent extreme conditions occur.
References


Morcrette, J. J. and P. Y. Deschamps, 1986: Downward longwave radiation at the surface in clear sky atmospheres: Comparison of measured,


Smith, W. L. and H. M. Woolf, 1983: Geostationary satellite sounder (VAS) observations of longwave radiation flux. paper presented at the Satellite System to Measure Radiation Budget Parameters and


