OBSERVATIONS OF THE MARTIAN ATMOSPHERE: THEMIS-VIS CALIBRATION, MESOSPHERIC CLOUDS, AND THE POLAR VORTEX

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Timothy Hubbard McConnochie

January 2007

OBSERVATIONS OF THE MARTIAN ATMOSPHERE: THEMIS-VIS CALIBRATION, MESOSPHERIC CLOUDS, AND THE POLAR VORTEX

Timothy Hubbard McConnochie, Ph.D.

Cornell University 2007

We present observations of the Martian atmosphere derived from two instruments: the Thermal Emission Spectrometer (TES) on the Mars Global Surveyor spacecraft, and the visible light subsystem of the Thermal Emission Imaging System (THEMIS-VIS) on the Mars Odyssey spacecraft.

For TES, we start with vertically resolved temperatures derived as described by Conrath et al. (2000, JGR, 105), and from them we derive horizontal winds and Ertel potential vorticity on a time series of regular three-dimensional grids. The Ertel potential vorticity is used as a dynamical tracer and diagnostic tool to study the behavior of the martian polar vortices. We find that, in contrast to the terrestrial polar vortices, the martian polar vortices' Ertel potential vorticity typically has an annular maximum well away from the pole. We also find that the martian northern winter vortex is better organized than the southern winter vortex, and thus is likely to be a more effective barrier to mixing.

For THEMIS-VIS we develop a complete radiometric calibration pipeline. This pipeline is used for standard data processing to convert Engineering Data Records (EDRs) to the Reduced Data Records (RDRs) released by NASAs Planetary Data System. We use THEMIS-VIS nadir-pointed images to detect clouds in the 40 km to 80 km altitude range, measuring altitude from parallax and velocity from

cross-track motion during the imaging sequence. We have observed 5 cases of aphelion season equatorial high-altitude clouds during late afternoon, all located in the eastern Tharsis / Valles Marineris region, and 30 cases of high-altitude cloud features in the northern winter (perihelion season) mid-latitudes, all but one in the Acidalia region. A simple radiative transfer model yields optical depths greater than 0.2 for the equatorial clouds, as well as constraints on their composition. The mid-latitude high-altitude features are visible only in twilight, a geometry for which our simple plane parallel radiative transfer model is not valid. Comparing the zonal velocity of the clouds with a radiative transfer model, we find good agreement in the northern winter mid-latitudes, but poorer agreement for equatorial clouds.

BIOGRAPHICAL SKETCH

Timothy McConnochie was born on September, 15, 1975, in Rochester, New York. He spent most of his childhood in nearby Pittsford, NY, where he graduated from Pittsford Sutherland High School in 1994. McConnochie received his BA from Williams College in Williamstown, MA, in 1998. Prior to enrolling at Cornell in August, 2000, he worked at the Board of Governors of the Federal Reserve System in Washington, DC, and at the Jet Propulsion Laboratory, in Pasadena, CA.

ACKNOWLEDGEMENTS

This work was supported in part by all of the following: a National Science Foundation Graduate Research Fellowship, NASA Mars Data Analysis Program grants NASG5-10580 and NNG04GJ10G, and by a grant to Professor Bell from the NASA Mars Odyssey Orbiter Participating Scientist program (JPL Contract 1241434).

I acknowledge all of the following people: Mike Caplinger of Malin Space Science Systems, for helping me to understand the THEMIS-VIS detector design; Steve Silverman and Hansford Cutlip of Raytheon Santa Barbara Remote Sensing for their help in gathering necessary records from the THEMIS-VIS pre-flight calibration activities; Jason Soderblom for providing photometric coefficients for the HST data, Alex Hayes and Min Hubbard for additional help with the HST data set; Kim Murray and Noel Gorelick of Arizona State University for their feedback on the results of the THEMIS-VIS calibration pipeline; Laural Cherednik, Andras Dombovari, and Kelly Bender, the THEMIS operations planners, for making sure that I got the data I needed; Phil Christensen, the THEMIS PI, for trusting me with the THEMIS-VIS calibration; Ken Herkenhoff and Paul Geissler for thoughtful and influential reviews of the THEMIS-VIS calibration manuscript; Andy Ingersoll and John Wilson, for valuable reviews of the Mars Polar Vortex manuscript; Josh Bandfield for sharing his work on TES error analysis; Huiqun Wang for sharing her MOC Daily Global Maps and her insights on them; Melissa Strausberg, for helping me to figure out who I needed to get in touch with; Barney Conrath, Don Banfield, and Mike Smith for their support, advice and patience with the Polar Vortex paper; Mark Richardson for getting the ball rolling on the THEMIS atmospheric campaign; Dmitry Savransky for keeping the computers running, and for reverse engineering ISIS, which made the mesospheric cloud mea-

surements possible; Anthony Toigo, for sharing his GCM results, and helping me to understand them; and Mike Wolff, for helping me to figure out how to do radiative transfer modeling, and for several years worth of patience and encouragement.

It goes without saying that I am intellectually indebted to my advisors: Jim Bell and Peter Gierasch. More importantly, I wish to thank them both for offering and endless supply of moral support, and patience, and accessibility. I cannot imagine a more ideal pair of mentors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The modern understanding of the Martian atmosphere has been comprehensively reviewed by Zurek [1992], and in companion papers by Owen [1992], James et al. [1992], Zurek et al. [1992], Jakosky and Haberle [1992], and Kahn et al. [1992]. (See Encrenaz [2001] for a more recent but much briefer overview.) As Zurek [1992] discusses, analogies between the terrestrial and martian atmospheres are an important tool in the study of both. Mars's atmosphere is uniquely analogous to Earth's, because it is a thin (scale height $<<$ radius) layer on a rapidly rotating planet with earth-like obliquity, and because its average temperature (~ 200 K) is, astronomically speaking, comparable. This work focusses on observations of two terrestrial analogs in the martian atmosphere: the winter polar vortex, using temperature sounding data from the Mars Global Surveyor Thermal Emission Spectrometer (MGS-TES) [Conrath et al., 2000]; and mesospheric clouds, using data from the Mars Odyssey Thermal Emission Imaging System Visible Imaging Subsystem (THEMIS-VIS). We also describe the radiometric calibration of THEMIS-VIS.

Beginning in 1997, both MGS-TES [*e.g.*, Conrath et al., 2000; Pearl et al., 2001; Smith et al., 2000, 2001; Smith, 2004] and the Mars Orbiter Camera Wide-Angle detrectors [*e.g.*, Cantor et al., 2002; Wang and Ingersoll, 2002; Benson et al., 2003; Wang et al., 2005] (also on MGS) have contributed to a comprehensive record of the seasonal behavior of the martian atmosphere, one which is still broadly consistent with the Zurek [1992] overview. Key factors which influence Mars's seasonal behavior in the present epoch include:

1. Typical surface pressures are $\sim$ 6 mb, with the consequence that the clear atmosphere has limited influence on the surface temperature.

2. Mars has an orbital eccentricity of 0.09, with perihelion occurring near northern winter solstice. Thus Mars receives $\sim$45% more solar flux in northern winter than in southern winter.

3. Polar night surface temperatures fall to the saturation point of CO_2, which makes up 95% of the atmosphere by volume, leading to a highly repeatable pattern of global pressure variations in which $\sim$30% of the mass of the atmosphere cycles into and out of the seasonal CO_2 polar caps.

4. Water ice in the southern polar cap is mostly covered by a veneer of CO_2 ice all year round, but in the northern polar cap it is exposed during the summer.

The large seasonal difference in flux makes it intuitive to describe the two extremes of the martian year as the "perihelion season" and the "aphelion season". The aphelion season is cooler, and characterized by an equatorial belt of water ice clouds [*e.g.*, Wolff et al., 1999], and by a large pulse of water vapor into the atmosphere which originates in the northern polar region as the H_2O cap is exposed. The mass of the atmosphere reaches its minimum late in the aphelion period due to the accumulation of CO_2 ice during the longer southern winter. [Zurek, 1992; Smith, 2004]

The perihelion season is characterized by higher temperatures, steeper temperature gradients, stronger winds, greatly decreased equatorial water ice aerosol opacity, and major, sometimes planet encircling, dust storms. Since the silicate dust absorbs a significant amount of longwave radiation, the perihelion period tem-

peratures and circulation are highly responsive to dust loading, creating a positive feedback which is presumably important in the development of major storms, and making perihelion period conditions highly variable [Smith, 2004] from year to year.

Even though the southern polar cap mostly retains its veneer of CO_2 ice, there is still a pulse of water vapor from the perihelion summer (southern) pole, but this pulse contains a much smaller amount of water vapor than does its aphelion period counterpart. Just as the perihelion summer pole releases less water vapor, the perihelion winter pole accumulates less CO_2. A secondary minimum in the mass of the atmosphere, caused by condensation and accumulation of CO_2 ice at the north pole, occurs late in the perihelion period, near the end of northern hemisphere winter, but based on the Viking Lander pressure record the mass in the atmosphere at this time is $\sim$20% greater than during the southern winter / aphelion period pressure minimum. [Zurek, 1992, and references therein]

Although equatorial condensate clouds are greatly diminished during the perihelion season, orographic clouds associated with the major shield volcanoes are present all year round. The other important group of condensate clouds is the polar hood, present at both poles during winter, but most visible in the north because its fringes extend beyond the boundary of polar night. The visible portions of the polar hood are water ice [James et al., 1992, and references therein], and the behavior of the north polar hood cloud as seen by MOC-WA is described by Wang and Ingersoll [2002]. However, the Mars Orbiter Laser Altimeter (MOLA) on MGS records clouds deep inside polar night which have been interpreted as CO_2 [Colaprete and Toon, 2002; Colaprete et al., 2003].

Important features of the Martian circulation that are evident from MGS-

TES and common to both seasons include a strong westerly winter polar jet (the polar vortex, which is the focus of Chapter 2), easterlies in the equatorial mesosphere, and maximum mesospheric temperatures that occur in the winter polar region. [Smith et al., 2001]. General circulation models [*e.g.*, Pollack et al., 1981; Haberle et al., 1993] indicate that heat transport at low to mid- latitudes is dominated by a single cross-equatorial Hadley cell during the solistices, and by a pair of Hadley cells at the equinoxes. The models also suggest that "condensation flow" [Pollack et al., 1981] dominates heat transport in winter high latitudes.

Both the winter polar vortex (Chapter 2) and mesospheric clouds (Chapter 3) are key indicators, and potentially drivers, of the martian climate. The polar vortex, as will be discussed, tends to isolate the winter pole, where the the crucial CO_2 condensation is taking place, from the rest of the atmosphere. Thus the extent of this isolation has implications for the amount of CO_2 sequestered at the poles, and so understanding the mechanism of the isolation might be important for understanding how the mass of the martian atmosphere is evolving and has evolved. Since water and possibly dust are also sequestered at the winter pole, the polar vortex is potentially important for the distribution of these constituents as well. The comparable dynamical structure on Earth, the stratospheric polar vortex, enables stratospheric ozone depletion by means of the isolation it causes [*e.g.*, Schoeberl and Hartmann, 1991], and it has been shown to respond to global climate trends [Thompson et al., 2000].

Mesospheric clouds are potentially indicators of gravity wave breaking and hence the wave drag which, as discussed by, *e.g.*, Joshi et al. [1995], is a crucial boundary condition for the martian general circulation. The aerosols themselves might also be important for the chemistry of the mesosphere: Atreya and Bla-

mont [1990] for example suggest that these aerosols could help recycle CO back to CO_2 in the martian mesosphere, thereby reducing the long term loss of CO_2 to space, with obvious implications for long term climate change. Martian mesospheric clouds and their terrestrial counterparts, noctilucent clouds, share similar problems in models of their formation mechanisms. For example, in both cases, heterogeneous nucleation is apparently required, but the source of the condensation nuclei is unclear. If studies of martian mesospheric clouds lead to insights into terrestrial noctilucent cloud formation, they may help to explain the interesting and potentially important terrestrial climate change phenomenon that reports of terrestrial noctilucent clouds have shown an upward secular trend [Klostermeyer, 2002] throughout the industrial era.

CHAPTER 2

OBSERVATIONS OF THE MARS POLAR VORTEX WITH THE MARS GLOBAL SURVEYOR THERMAL EMISSION SPECTROMETER

2.1 Introduction

The winter season, westerly circumpolar flow of the martian atmosphere, like that of the terrestrial stratosphere, is concentrated into a jet which lies near 60 degrees latitude. This jet is known as the polar vortex. Similar dynamical structures have been observed on Jupiter [Orton et al., 2002] and Venus [Taylor, 2002]. Polar vortices are of interest because they act as a barrier, inhibiting energy transport and potentially preventing the mixing of aerosols and chemical species. Thus, they control the response of winter polar processes to climatic forcings, both short term and long term.

The terrestrial polar vortex has been understood to be the dynamical controlling mechanism for ozone depletion in the polar stratosphere [*e.g.*, Schoeberl and Hartmann, 1991] for more than a decade. An example of the importance of dynamical control is that on Earth the southern hemisphere experiences greater ozone depletion than the north, because the southern winter polar vortex is a more effective barrier to mixing [Schoeberl et al., 1992; Andrews et al., 1987]. More recently, the Earth's stratospheric annular modes, which are essentially a weakening/strengthening oscillation of the polar vortex jet, have been shown to be coupled to, and possibly even a driving mechanism for, the tropospheric Arctic Oscillation (AO) / North Atlantic Oscillation (NAO) phenomenon [Baldwin and Dunkerton, 2001]. The AO / NAO is a key player in northern temperate zone win-

ter weather patterns [Thompson and Wallace, 2000] and an increasing bias towards the positive (stronger polar vortex) phase in the past 30 years has been shown to be correlated with global warming trends [Thompson et al., 2000].

The terrestrial polar vortex exhibits a dramatic phenomenon known as a "sudden stratospheric warming." A typical sudden stratospheric warming is characterized by an increase of 40 - 60 K in average polar temperatures within a one week period. In major midwinter stratospheric warming events, which are observed primarily in the northern hemisphere and with frequencies of about once every two years, the circumpolar flow actually reverses, becoming easterly [Andrews et al., 1987]. Early interest in the properties of the Mars polar vortex was triggered by Viking Orbiter observations of a similarly abrupt 60 K warming in polar temperatures during the northern hemisphere winter of 1977. The warming event occurred during the second of two global dust storms that occurred that season [Martin and Kieffer, 1979; Jakosky and Martin, 1987].

Barnes and Hollingsworth [1987] suggested that this warming event on Mars was in fact analogous to the terrestrial stratospheric warmings. However, GCM simulations by Wilson [1997] suggest that the martian polar warming could have been basically symmetric about the pole, associated with an expanded Hadley circulation. In Wilson's simulations, the expanded Hadley cell was enabled by momentum flux divergence attributed to dust-enhanced diurnal tides. In contrast, terrestrial stratospheric warmings are caused by breaking planetary-scale Rossby waves [*e.g.*, McIntyre and Palmer, 1983], and hence are highly asymmetric about the pole. Unfortunately, the Viking measurements include only one longitude position per sol [Jakosky and Martin, 1987], making it difficult to measure vortex asymmetries or evaluate the importance of planetary waves.

The Mars Global Surveyor (MGS) Thermal Emission Spectrometer (TES) data set [Christensen et al., 1998] provides over three martian years of consistently sampled, vertically resolved temperature measurements [Conrath et al., 2000; Smith et al., 2001], allowing us to take a comprehensive look at the range of behaviors exhibited by the Mars polar vortex. Banfield et al. [2000, 2003, 2004] and Wilson et al. [2002] have used these data to detect forced and traveling wave modes in the martian atmosphere. Of particular relevance for the polar vortex, Banfield et al. show that the polar jet supports stationary [Banfield et al., 2003] and traveling [Banfield et al., 2004] waves of zonal wave numbers 1, 2 and 3, with most of the amplitude in zonal wave number 1. They also show that there exists a zone of symmetric instability on the equatorward flank of the polar jet, as well as regions on both the equatorward and poleward flanks of the jet where the barotropic instability criterion is satisfied [Banfield et al., 2004].

The analysis quantity of choice for studies of the terrestrial polar vortex is Ertel potential vorticity, *PV* [*e.g.*, Schoeberl and Hartmann, 1991; Nash et al., 1996; Lait, 1994; Waugh and Randel, 1999]. *PV*, in the absence of diabatic (friction and heating) effects, is a conserved tracer of atmospheric motion. On Earth, *PV* gradients are correlated with gradients in the concentrations of chemical species [Schoeberl et al., 1992], and *PV* provides a convenient diagnostic for events, such as planetary wave breaking, that mix air across the vortex boundary [McIntyre and Palmer, 1983]. Since potential temperature, θ, is also conserved in the absence of diabatic effects, contours of constant *PV* on a θ surface represent actual physical barriers to mixing on dynamical timescales. Flow across *PV* contours *will* occur on radiative timescales. For example, Nash et al. [1996] shows that the early winter increase of *PV* in the polar stratosphere is accomplished by radiatively induced circulation. An

additional advantage of *PV* is its effect on planetary wave propogation . Vorticity is essentially the working substance for Rossby waves, and therefore regions of steep vorticity gradient, like the polar vortex edge, can act to refract and trap planetary waves [Polvani and Saravanan, 2000]. *PV* maps also highlight regions of instability with respect to wave motion, as discussed by Banfield et al. [2004].

In light of these advantages, and for the sake of direct planet-to-planet comparison, we have adopted *PV* as our primary analysis quantity for the Mars polar vortex. We convert the MGS-TES temperature retrievals into a uniform longitude-latitude-altitude-time grid, and then use this gridded data set to calculate *PV*. We must, of course, be wary of the fact that radiative time scales on Mars are generally shorter than on Earth, so that the time scale over which PV is conserved could be much shorter in certain regions, depending on the degree of departure from radiative equilibrium.

Another concern, unique to Mars, is that CO_2 condensation is a dynamically important process [Pollack et al., 1990]. Recent observations suggest that CO_2 condensate clouds are pervasive inside the polar vortex [Ivanov and Muhleman, 2001; Colaprete and Toon, 2002; Colaprete et al., 2003]. For this reason, we also consider gridded estimates of temperature, focusing in particular on the difference between polar temperatures and the CO_2 saturation point.

The objectives of this paper are primarily descriptive. We begin by presenting the methods used to derive gridded temperatures, winds, and *PV*, and then present a series of maps and cross sections to describe the spatial pattern, temporal evolution, and inter-annual variability of these quantities for both the north and south polar vortices. We pay particular attention to two polar warming events — one in southern winter, and one in northern winter. Our discussion addresses those

polar vortex behaviors which we believe to be most in need of further investigation, focussing on pole-to-mid-latitude mixing, the influence of dust, and the causes of the polar warmings.

2.2 Methodology

2.2.1 TES temperature retrievals

We start with MGS-TES retrievals of temperature as a function of pressure. The derivation of these temperature profiles from the 15 μm CO_2 absorption profile is discussed by Conrath et al. [2000]. In consideration of the intrinsically limited resolution of these measurements, and in order to prevent our working data set from becoming unmanageably large, we sample the derived temperature profiles at intervals of one-half scale height. TES nadir-pointed observations have denser sampling in the horizontal dimensions than TES limb-pointed observations. For this reason, we have used only nadir-pointed observations for the work presented here. As a result, our data set is limited to the lowest four scale heights of the atmosphere.

For the mapping phase of its mission, MGS has maintained a sun synchronous polar orbit such that nadir-pointed TES observations always occur at roughly 2 PM local solar time on the day side of the orbit, and 2 AM local solar time on the night the of the orbit. The precise solar time of TES observations is actually a function of latitude, and at any given latitude, the local solar time of TES observations varies, primarily as a function of season (see Fig. 2.1). MGS's 2-hour orbital period causes the ground tracks of individual orbits to be separated by approximately 30 degrees of longitude. Temperature profiles are obtained along the MGS ground track at

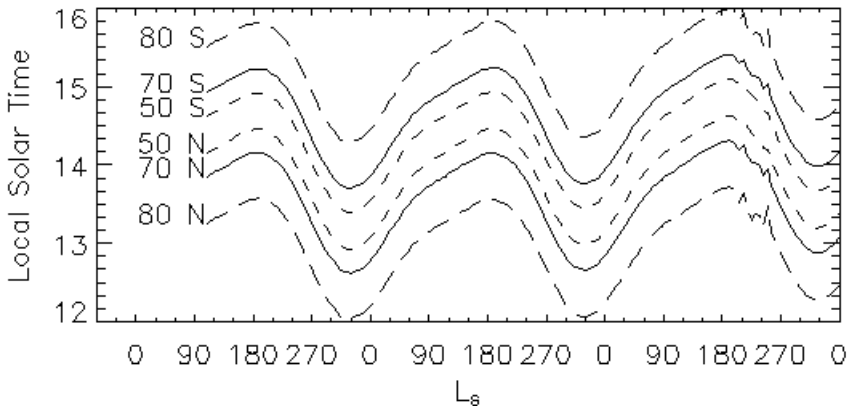

Figure 2.1: Local solar time for three years of TES mapping at selected latitudes.

typical intervals of 0.1° in latitude. Gaps in this sampling sequence are common. The pattern of temperature measurement locations produced by a typical sol of sampling is shown in Fig. 2.2. Banfield et al. [2003] provide further discussion of the capabilities and limitations of this data set. The temperature retrievals and vertical sampling thereof used here are identical to that of Banfield et al. [2003].

2.2.2 Temperature analysis grid

In order to support the calculation of dynamically important quantities such as winds and potential vorticity, we interpolate the TES temperature retrievals onto a time series of grids with uniform sampling in latitude, longitude, and altitude in log-pressure coordinates. All of the temperatures in a given grid are for the same instant in time, meaning that we additionally interpolate in the temporal dimension. We generate one set of grids for the daytime temperatures, and another set for the nighttime temperatures. Since all of the source measurements are for ∼2 PM local solar time (or ∼2 AM for the nighttime grids), each grid obviously contains none of the diurnal variation that would exist in a true instantaneous measurement of global temperatures. In other words, our analysis grids are maps

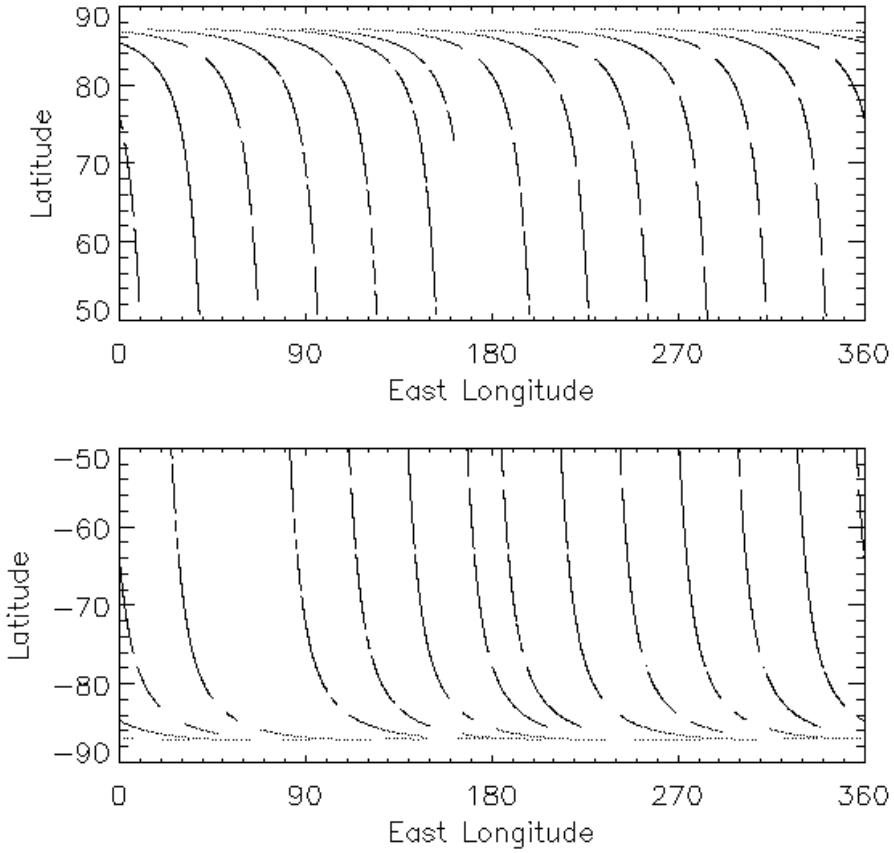

Figure 2.2: Location of TES nadir-pointed temperature measurements for a typical one-sol mapping period in the northern (top) and southern (bottom) polar regions. Each measurement location is plotted as a single point, but these points are generally so close together that they appear as a continuous line in this figure. Gaps in this line indicate pauses in the normal sampling sequence.

of instantaneous daytime (or nighttime) temperatures with the diurnal component "removed."

Table 2.1 summarizes the analysis grid parameters. The temperature analysis grids are generated as follows:

1. Classify each TES retrieval as either "daytime" or "nighttime." "Daytime" retrievals are defined as those with local solar times of more than 6 and less than 18 hours. "Daytime" retrievals are used to populate the daytime grids, and "nighttime" retrievals are used to populate the nighttime grids.

2. Sample the TES temperature retrievals at intervals of one-half scale height, starting from the 6.1 millibar pressure level. In the remaining steps, each pressure level is handled separately.

3. For each orbit of data points, establish a uniform latitude grid by smoothing, resampling, and filling in data gaps, as follows:

   (a) At each grid latitude, ϕ_j, the temperature, T_j, longitude, λ_j, and time, t_j, are determined from a weighted average of neighboring data point temperatures, T_n, longitudes, λ_n, and times, t_n:

$$\left.\begin{aligned} T_j &= \frac{\sum_n W_n T_n}{\sum_n W_n} \\ \lambda_j &= \frac{\sum_n W_n \lambda_n}{\sum_n W_n} \\ t_j &= \frac{\sum_n W_n t_n}{\sum_n W_n} \end{aligned}\right\} \quad \text{if and only if} \quad \sum_n W_n \geq minWeight \tag{2.1}$$

   where $minWeight$ is as given in Table 2.1, and the sum is over all data points in an orbit. Grid points for which $\sum_n W_n < minWeight$ are treated as missing. The weight W_n of the n-th data point at a given

grid latitude is a function of the latitude of the data point, ϕ_n:

$$W_n = \begin{cases} \mathrm{e}^{-\left(\frac{\phi_j - \phi_n}{s}\right)^2} & \text{if } \left|\frac{\phi_j - \phi_n}{s}\right| \leq 3 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

where s is given in Table 2.1.

(b) Gaps between valid grid points, *i.e.*, clusters of grid points which are "missing" because $\sum_n W_n$ is too small, are filled in if the the gap size, $\phi_{j_{\text{high}}} - \phi_{j_{\text{low}}}$, where $\phi_{j_{\text{high}}}$ and $\phi_{j_{\text{low}}}$ are the latitudes of the valid grid points on either side of the gaps, is less than the parameter *maxLatGap* (see Table 2.1). Gap filling is accomplished with a simple linear interpolation from the j_{high} and j_{low} grid points. Explicity:

$$\text{if} \quad \left(\sum_n W_n\right)_j < minWeight \text{ and } \left|\phi_{j_{\text{high}}} - \phi_{j_{\text{low}}}\right| \leq maxLatGap$$

$$\text{then} \quad \begin{cases} T_j = \frac{\left|\phi_{j_{\text{high}}} - \phi_j\right| T_{j_{\text{high}}} + \left|\phi_{j_{\text{low}}} - \phi_j\right| T_{j_{\text{low}}}}{\left|\phi_{j_{\text{high}}} - \phi_{j_{\text{low}}}\right|} \\ \lambda_j = \frac{\left|\phi_{j_{\text{high}}} - \phi_j\right| \lambda_{j_{\text{high}}} + \left|\phi_{j_{\text{low}}} - \phi_j\right| \lambda_{j_{\text{low}}}}{\left|\phi_{j_{\text{high}}} - \phi_{j_{\text{low}}}\right|} \\ t_j = \frac{\left|\phi_{j_{\text{high}}} - \phi_j\right| t_{j_{\text{high}}} + \left|\phi_{j_{\text{low}}} - \phi_j\right| t_{j_{\text{low}}}}{\left|\phi_{j_{\text{high}}} - \phi_{j_{\text{low}}}\right|} \end{cases} \tag{2.3}$$

4. At each grid pressure level p_k , and at each grid latitude ϕ_j, a bilinear interpolation of the T_j temperatures in all of the available orbits is performed to map them from the λ_j longitudes and t_j times onto the grid longitudes λ_i and grid times t_l, yielding a time series of gridded temperatures, T_{ijkl}. If the gap in longitude or time between adjacent non-missing λ_j or t_j points is larger the *maxLonGap* or *maxTimeGap* parameters, respectively, then all of the T_{ijkl} values that fall between these points are considered to be missing.

Table 2.1: Analysis parameters

Parameter	Value	Units	Description
Δ_λ	10.0	° longitude	grid spacing in longitude
Δ_ϕ	1.0	° latitude	grid spacing in latitude
$\Delta_{\ln p}$	0.5	–	vertical grid spacing, uniform in log-pressure coordinates
Δ_t	1.0	sols	grid spacing in time (1 sol = 88775 seconds)
p_0	6.1	millibars	lowest pressure level in the analysis grid
$p_{\max}$	$6.1 \times \mathrm{e}^{-7\Delta_{\ln p}}$	millibars	highest pressure level in the analysis grid
minWeight	4.5	–	minimum weight for temperature smoothing (section 2.2.2)
s	1.0	° latitude	temperature retrieval smoothing parameter (section 2.2.2)
maxLatGap	6.0	° latitude	maximum latitude gap that is filled by interpolation (section 2.2.2)
maxLonGap	70.0	° longitude	maximum longitude gap that is filled by interpolation (section 2.2.2)
maxTimeGap	1.5	sols	maximum time gap that is filled by interpolation (section 2.2.2)
k_{ref}	1	–	vertical grid index of the reference pressure level for geopotential height calculations (section 2.2.3)
$p(k_{\mathrm{ref}})$	3.7	millibars	reference pressure level for geopotential height calculations (section 2.2.3)
smoothLat	5.0	° latitude	wind solution latitude smoothing parameter (section 2.2.4)
smoothLon	50.0	° longitude	wind solution longitude smoothing parameter (section 2.2.4)
scaleFactor	0.25	–	controls the rate at which the wind solution adjusts to the residuals at each iteration (section 2.2.4)

Table 2.2: Adopted Physical Constants

Parameter	Value	Units	Description	Reference
R	192.	$\mathrm{Joules\,kg^{-1}\,K^{-1}}$	Specific gas constant for the martian atmosphere.	Zurek et al. [1992]
g_0	3.72	$\mathrm{m\,s^{-2}}$	Surface gravity	Zurek et al. [1992]
Ω	7.088×10^{-5}	$\mathrm{s^{-1}}$	angular velocity of Mars	Zurek et al. [1992]
a	3394.	km	radius of Mars	Zurek et al. [1992]
c_p	734.	$\mathrm{Joules\,kg^{-1}\,K^{-1}}$	specific heat at constant pressure (assumed constant)	Magalhães et al. [1999]

2.2.3 Geopotential Height

We assume hydrostatic equlibrium in order to calculate the geopotential height, Z_{ijkl}, above some reference pressure level, $p = p(k_{\mathrm{ref}})$, by vertical integration of the temperature field.

$$Z_{ijkl} = \frac{\Delta_{\ln p}}{2} \frac{R}{g_0} \sum_{k'=k_{\mathrm{ref}}+1}^{k} \left(T_{ijk'l} + T_{ij(k'-1)l} \right) \tag{2.4}$$

R is the specific gas constant, and g_0 is the surface gravity (see Table 2.2). $\Delta_{\ln p}$ is the vertical grid spacing (which is a constant in log-pressure coordinates — see Table 2.1). We treat the geopotential height as a constant at $k = k_{\mathrm{ref}}$, and arbitrarily define $Z_{ijk_{\mathrm{ref}}l} = 0$. Since the main purpose of our geopotential height calculations is to facilitate subsequent calculation of winds, this constant geopential assumption is essentially equivalent to the usual assumption of zero wind at some reference pressure level. We have chosen $k_{\mathrm{ref}} = 1$. Given our grid specifications (Table 2.1), $p(k_{\mathrm{ref}}) = 3.70\,\mathrm{mb}$.

If, at a certain i, j, and l, a $T_{k_{\mathrm{ref}}}$ is missing, then for purposes of calculating Z_k, $T_{k_{\mathrm{ref}}}$ is found by linear extrapolation of T from the two lowest non-missing T_k. This extrapolation is necessary to prevent persistent data gaps at or near the locations of topographic highs. If there are fewer than 2 points at the given i, j, and l with $k \geq k_{\mathrm{ref}}$, then all of the Z_k at that i, j, and l are considered to be "missing."

2.2.4 Non-linear balance winds

We calculate horizontal winds from the geopential height data using a variation of the non-linear balance wind method suggested by Randel [1987]. This method seeks an approximate solution to the full primitive equations for zonal and merid-

ional momentum balance, neglecting only the local time derivatives ($\frac{\partial}{\partial t}$) and the vertical wind components. We generate wind estimates for latitudes poleward of $\pm 35^\circ$. The calculation is performed separately for each pole.

The balance equations are:

$$\epsilon_u = u + \frac{g_0}{af}\frac{\partial Z}{\partial \phi} + \frac{1}{af}v\frac{\partial v}{\partial \phi} + \frac{\tan\phi}{af}u^2 + \frac{1}{af\cos\phi}u\frac{\partial v}{\partial \lambda} \quad (2.5)$$

$$\epsilon_v = v - \frac{g_0}{af\cos\phi}\frac{\partial Z}{\partial \lambda} - \frac{1}{af\cos\phi}u\frac{\partial u}{\partial \lambda} + \frac{\tan\phi}{af}uv - \frac{1}{af}v\frac{\partial u}{\partial \phi} \quad (2.6)$$

where f is the Coriolis parameter, $f = 2\Omega\sin\phi$, a is the radius of Mars, Ω is the angular velocity of Mars about its rotation axis, u is the eastward wind component, and v is the northward wind component. An exact solution to the balance equations would have $\epsilon_u = 0$ and $\epsilon_v = 0$. However, an exact solution does not in general exist, and we therefore seek an approximate solution for the u and v fields that minimizes the ϵ_u and ϵ_v fields in a root-mean-square sense.

To evaluate partial derivatives at each grid point, we use centered derivatives wherever feasible. If either of the two neighboring grid points required for a centered derivative do not exist (= DNE), because of "missing" data or because it is off the edge of the grid, then the partial derivative is calculated using the data value at the grid point together with whichever neighboring value is available. If the data value AT the grid point itself is "missing," then the derivative at that grid point will considered "missing" as well, regardless of the validity of the neighboring points. To summarize, the partial derivatives for a certain quantity x along the "i-axis" of the grid are calculated as follows: (The j, k, and l subscripts are suppressed for clarity.)

$$\text{if } (x_i = \text{DNE}) \text{ or } [(x_{i-1} = \text{DNE}) \text{ and } (x_{i+1} = \text{DNE})] \quad \text{then } \left(\frac{\partial x}{\partial \lambda}\right)_i = \text{DNE} \quad (2.7a)$$

otherwise

$$\text{if } (x_{i-1} = \text{DNE}) \text{ and } (x_{i+1} \neq \text{DNE}) \qquad \text{then } \left(\frac{\partial x}{\partial \lambda}\right)_i = \frac{x_{i+1} - x_i}{\Delta_\lambda} \tag{2.7b}$$

$$\text{if } (x_{i-1} \neq \text{DNE}) \text{ and } (x_{i+1} = \text{DNE}) \qquad \text{then } \left(\frac{\partial x}{\partial \lambda}\right)_i = \frac{x_i - x_{i-1}}{\Delta_\lambda} \tag{2.7c}$$

$$\text{if } (x_{i-1} \neq \text{DNE}) \text{ and } (x_{i+1} \neq \text{DNE}) \qquad \text{then } \left(\frac{\partial x}{\partial \lambda}\right)_i = \frac{x_{i+1} - x_{i-1}}{2\Delta_\lambda} \tag{2.7d}$$

where Δ_λ is longitudinal grid spacing. Partial derivatives in the other dimensions are defined analogously. Note that this method of calculating derivatives makes it unnecessary to explicitly specify boundary conditions for our wind solution.

In order to converge on a solution for u and v, we found it necessary to include smoothing in the horizontal spatial dimensions as part of the algorithm. This is accomplished by convolution with a gaussian kernel, S:

$$S_{i'j'} = \begin{cases} \mathrm{e}^{-\left(\frac{i'^2}{2{\sigma_i}^2}\right)} \cdot \mathrm{e}^{-\left(\frac{j'^2}{2{\sigma_j}^2}\right)} & \text{if } -3\sigma \leq i' \leq 3\sigma \text{ and } -3\sigma \leq j' \leq 3\sigma \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

where σ_i and σ_j are chosen such that the full-width-half-max (FWHM) of the gaussian in that dimension is equal to the *smoothLon* and *smoothLat* parameters, respectively (see Table 2.1). Missing or out-of-grid data values are simply ignored during the smoothing process. Any data point labeled "missing" before the smoothing will still be considered "missing" afterwards.

Prior to the start of the wind-finding algorithm, the geopotential height grid at each pressure level is smoothed according to Eq. 2.8, and the smoothed Z values are used at all stages of the algorithm. All of the operations involved in the algorithm are performed simultaneously on each pressure level, so that the final result at

each level is based on the same number of iterations. The wind solution is derived separately for each time step.

The initial guess for the wind solution is the gradient wind for the u component,

$$u = -\Omega a \cos\phi + \sqrt{(\Omega a \cos\phi)^2 - \frac{g_0}{\tan\phi}\frac{\partial Z}{\partial \phi}} \tag{2.9}$$

and geostrophic for the v component,

$$v = \frac{g_0}{af\cos\phi}\frac{\partial Z}{\partial \lambda} \tag{2.10}$$

At each iteration, the $(\epsilon_u)_m$ and $(\epsilon_v)_m$ residuals are evaluated by applying Eqs. 2.5 and 2.6 to the current wind-field estimate, (u_m, v_m). The halting criteria (described below) are then checked, and if they are not satisfied, the new wind estimate is calculated as follows:

$$\begin{aligned} u_{m+1} &= u_m - \text{Smooth}(\textit{scaleFactor} \cdot (\epsilon_u)_m) \\ v_{m+1} &= v_m - \text{Smooth}(\textit{scaleFactor} \cdot (\epsilon_v)_m) \end{aligned} \tag{2.11}$$

The "Smooth" function indicates smoothing by convolution with a gaussian kernel (Eq. 2.8) as previously described. We have used a *scaleFactor* of 0.25, because we have found that this value leads to a final wind solution with lower root-mean-squared (RMS) residuals.

We evaluate the quality of the wind solution at each iteration by measuring the RMS ϵ_u and RMS ϵ_v for each row of the analysis grid. (A row contains all of the grid points at a given latitude, pressure level, and time.) For purposes of determining when to halt the algorithm, this quality information is distilled to a single number, Q, which we define as the mean of the RMS ϵ_u for all rows at the 0.304 mb level with latitude between $\pm 40°$ and $\pm 85°$. Once the algorithm is halted, the final wind solution is the solution with the lowest value of Q. The algorithm halts whenever $Q_{m+1} > Q_m$, or when the total number of iterations reaches 10.

Figure 2.3: Geopotential height in km (left), eastward wind u in m/s (center), and northward wind v in m/s (right) at the 0.83mb pressure level for a sol in early northern winter. The map projection is polar stereographic, covering latitudes north of 40 degrees.

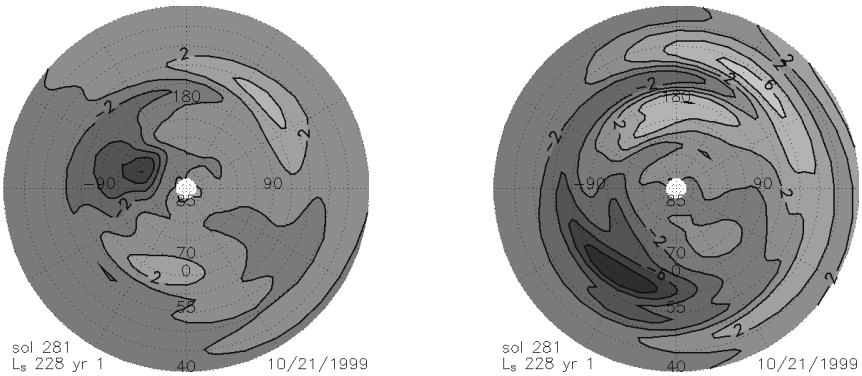

Figure 2.4: ϵ_u (left) and ϵ_v (right) residuals for the wind solutions shown in Fig. 2.3.

Figure 2.3 shows the geopotential height field, calculated u-component winds, and calculated v-component winds for a sol on which one of the highest-amplitude transient planetary waves was observed. Figure 2.4 shows ϵ_u and ϵ_v residuals for the same sol. These figures show that, in this particularly extreme circumstance, the balance residuals have an amplitude of about 25% of that of the calculated winds for the v-component. The balance residuals for the u-component winds are everywhere less than 10% of the calculated wind.

Figure 2.5, and Figs. 2.6 – 2.8 show the effect of other sources of uncertainty

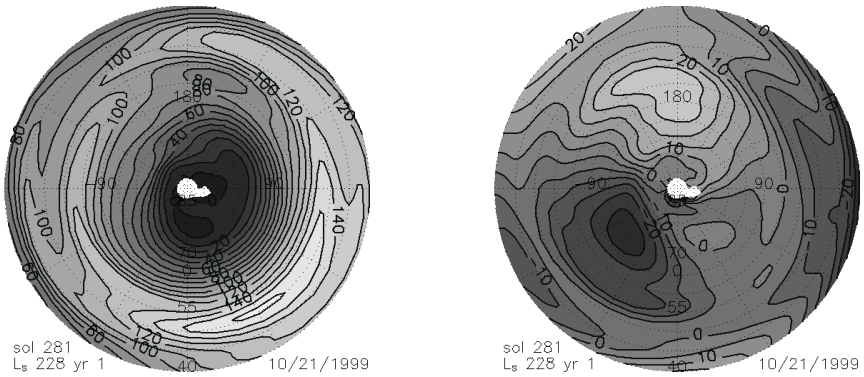

Figure 2.5: u (left) and v (right) winds as in Fig. 2.3, but calculated with a 9.50 mb lower boundary.

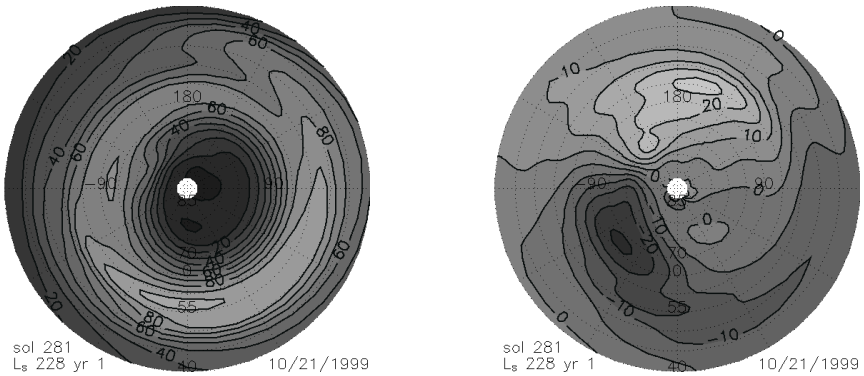

Figure 2.6: u (left) and v (right) winds as in Fig. 2.3, but calculated after the temperature data set has been perturbed by a global systematic error of 1.2 Watt $\mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,(\mathrm{cm}^{-1})^{-1}$ as described in the text.

in our wind solution. For the calculated winds, the most significant source of uncertainty is the choice of the lower boundary on which to assume zero wind. Figure 2.5 shows the effect of moving this assumed lower boundary to 9.50 mb. Doing so mainly affects the amplitude of the u-component winds, increasing them by $\sim$50%.

The other sources of error involve uncertainties in the TES radiance. Figures 2.6 – 2.8 show experiments with three types of radiance errors. To estimate the effect of a radiance error on the retrieved temperatures, we treat the nominal temperature

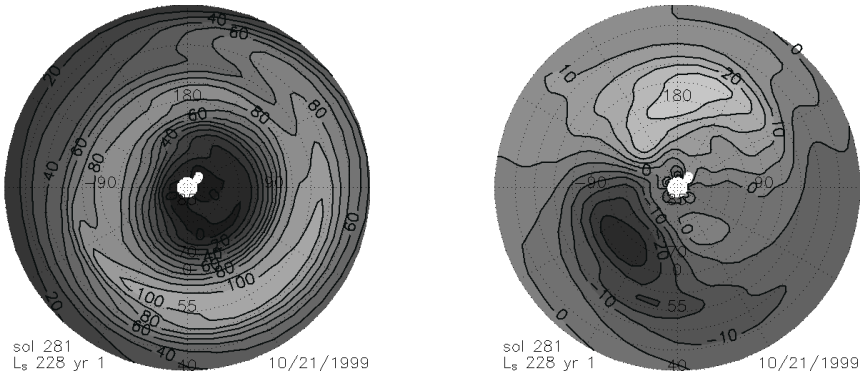

Figure 2.7: u (left) and v (right) winds as in Fig. 2.3, but calculated after individual measurements in the temperature data set have been perturbed by exaggerated random errors (2.0 Watt $\mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,(\mathrm{cm}^{-1})^{-1}$ RMS) as described in the text.

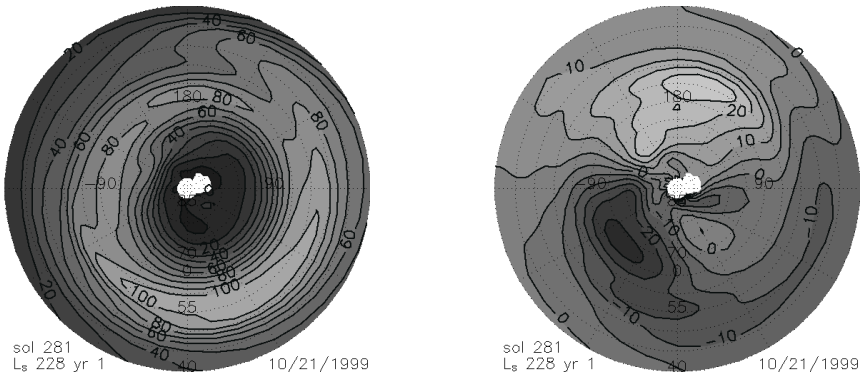

Figure 2.8: u (left) and v (right) winds as in Fig. 2.3, but calculated after each orbit of the temperature data set has been perturbed by 0.5 Watt $\mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,(\mathrm{cm}^{-1})^{-1}$ RMS random error (see text).

as a 667cm^{-1} (15 μm) brightness temperature, convert it to radiance, apply the radiance perturbation, and then convert back to temperature. Roughly speaking, for a given radiance perturbation, the temperature perturbation scales as T^{-4}. This means that radiance errors affect the temperatures gradients as well as absolute values, and so can influence the wind solution.

Fig. 2.6 shows wind results after adding 1.2 Watt cm^{-2} sr^{-1} (cm^{-1})$^{-1}$ to all retrievals, simulating a global systematic radiance error. This radiance perturbation is equivalent to +5 K at 145 K, and thus is similar to the systematic errors that Conrath et al. [2000] suggested may exist in the TES calibration. The perturbed wind results are only slightly ($\sim 10\%$) less than the preferred solution shown in Fig. 2.3.

Figure 2.7 shows wind results after adding random perturbations with an RMS of 2 Watt cm^{-2} sr^{-1} (cm^{-1})$^{-1}$ to the retrievals, simulating noise in the TES radiance data set . This noise levels is equivalent to $\pm$9 K at 145 K, much larger than the 1 K RMS errors suggested by Conrath et al. [2000]. Nevertheless, the wind solution is not noticeably altered, implying that random noise is not significant as a source of error.

The TES calibration relies on a series of space-looks acquired once each orbit to subtract instrumental background radiance. Noise in these space-look measurements propagates into an error which is random from orbit to orbit, but constant within an orbit (J. Bandfield, personal communication). In Fig. 2.8 we explore the effects of such errors by adding a random radiance perturbation to the retrievals from each orbit. We have used perturbations with an RMS of 0.5 Watt cm^{-2} sr^{-1} (cm^{-1})$^{-1}$ (equivalent to $\pm$2.2 K at 145 K). As the figure shows, the wind solution is not obviously altered by this perturbation, although we will

see in 2.2.7 that the effect *is* evident in the *PV* field.

2.2.5 Potential Temperature, θ

We define θ relative to the 6.1mb pressure level on Mars. We assume that c_p (see Table 2.2) is approximately constant for the lower atmosphere of Mars, adopting a value appropriate to the mean Mars atmospheric temperature of 200 K.

$$\theta = T\left(\frac{6.1\text{mb}}{p}\right)^{\frac{R}{c_p}} \tag{2.12}$$

2.2.6 Interpolation to θ vertical coordinate.

In order to show *PV* on θ coordinate surfaces, we simply apply a linear interpolation from the *PV* on log-pressure surfaces, calculated as described below, onto the desired θ coordinate surface using the θ values calculated for each grid point.

2.2.7 Ertel Potential Vorticity

Ertel potential vorticity, *PV*, is defined in an intertial frame as:

$$PV = \frac{(\vec{\nabla} \times \vec{v}) \cdot \vec{\nabla}\theta}{\rho} \tag{2.13}$$

where $\vec{v}$ is the velocity vector, and ρ is the density.

We calculate *PV* in the co-rotating, log-pressure coordinate system of our analysis grid as follows:

$$PV = \frac{g_0 \cos\phi \left(\frac{\partial\theta}{\partial\phi}\frac{\partial M}{\partial\hat{z}} - \frac{\partial\theta}{\partial\hat{z}}\frac{\partial M}{\partial\phi}\right) + g_0 \left(\frac{\partial\theta}{\partial\hat{z}}\frac{\partial N}{\partial\lambda} - \frac{\partial\theta}{\partial\lambda}\frac{\partial N}{\partial\hat{z}}\right)}{pa^2\cos^2\phi} \tag{2.14}$$

where

$$M = au\cos\phi + \Omega a^2\cos^2\phi \tag{2.15}$$

$$N = av\cos\phi \tag{2.16}$$

$\hat{z}$ is the log-pressure vertical coordinate.

Vertical derivatives are calculated as described in Eqs. 2.7a - 2.7d (with $\hat{z}$ and k replacing λ and i, respectively), except that all derivatives at a given grid pressure level are always calculated by the same method, *i.e.*, one and only one of Eqs. 2.7b - 2.7d. If more than two-thirds of derivatives calculated according to Eq. 2.7d would be "missing" (because one of the neighboring data points used to calculate it is missing), then all derivatives are calculated according to whichever of Eq. 2.7b or Eq. 2.7c would give the the largest number of non-missing values.

Figure 2.9a shows *PV* calculated for the same sol as in Figs. 2.3 – 2.8. In order to estimate the impact of the wind residuals on the potential vorticity, we have also plotted, in Fig. 2.9b, potential vorticity calculated using wind components that have been perturbed by twice the residuals of the final wind solution. In other words, Fig. 2.9b shows *PV* calculated using $u + 2\epsilon_u$ and $v + 2\epsilon_v$ for the wind components. Since the *PV* shown in Fig. 2.9b is essentially the same as in Fig. 2.9a, we conclude that errors in our wind solution are not a significant source of uncertainty in *PV*.

Figure 2.9c shows *PV* calculated for the alternative 9.50 mb lower boundary. Evidently, the choice of lower boundary has only a small effect on the magnitude of *PV* near the winter pole. The lower boundary does become significant further away from the pole, where the $\theta = 240$K surface is very close to the lower boundary.

Next, although it is not the goal of this paper to assess *diurnal* variations in the polar vortex, we do, in Fig. 2.9d, test the influence of time-of-day on the *PV*-field. This test is complicated by the fact that the nighttime data for a given longitude is acquired twelve hours earlier (and later) than the daytime data. Although our interpolation procedure attempts to correct for this effect, it is not entirely obvious

whether nighttime data from the *same* sol, or nighttime data from the *following* sol is a more appropriate comparison. We have solved this problem empirically by noting that the spiral structure of sol 281 daytime has not yet developed on sol 281 nighttime, but that the night of 282 is in fact quite similar to the day of 281. (Sol numbers are defined in Section 2.3.) We have therefore chosen to plot the night of 282 as Fig. 2.9e to illustrate that time-of-day does not have a large effect on the size, magnitude, and shape of the high-*PV* region.

Figures 2.9e–g illustrate the effects of TES radiance errors on the *PV*-field, just as Figs. 2.6 – 2.8 did for the derived wind fields. Figures 2.9e, 2.9f, and 2.9g show the results of global, random, and random-by-orbit errors, respectively, with the same magnitudes as in Figs. 2.6, 2.7, and 2.8, respectively. In Fig. 2.9e we see that the global perturbation leads to a substantial (about 25%) decrease in the magnitude of the calculated *PV*, but this hypothetical correction has no qualitative effect on the *PV* field — the overall shape, and the location of the maxima, for example, are unchanged. In Fig. 2.9f we find that even exaggerated random errors produce only minor changes in the shape of the *PV* field, indicating that random noise does not significantly contaminate our results. However, as Fig. 2.9g shows, our *PV* field is more sensitive to orbit-by-orbit random errors. For moderate $0.5\,\mathrm{Watt\,cm^{-2}\,sr^{-1}\,(cm^{-1})^{-1}}$ orbit-by-orbit pertubations, we notice additional longitudinal variations in the *PV* field, although the overall pattern remains intact. This means that for any given *PV* map, we expect that *some* of the longitudinal *PV* variations are the result of TES instrumental noise, but that major *PV*-field features like the sol 281 spiral are robust.

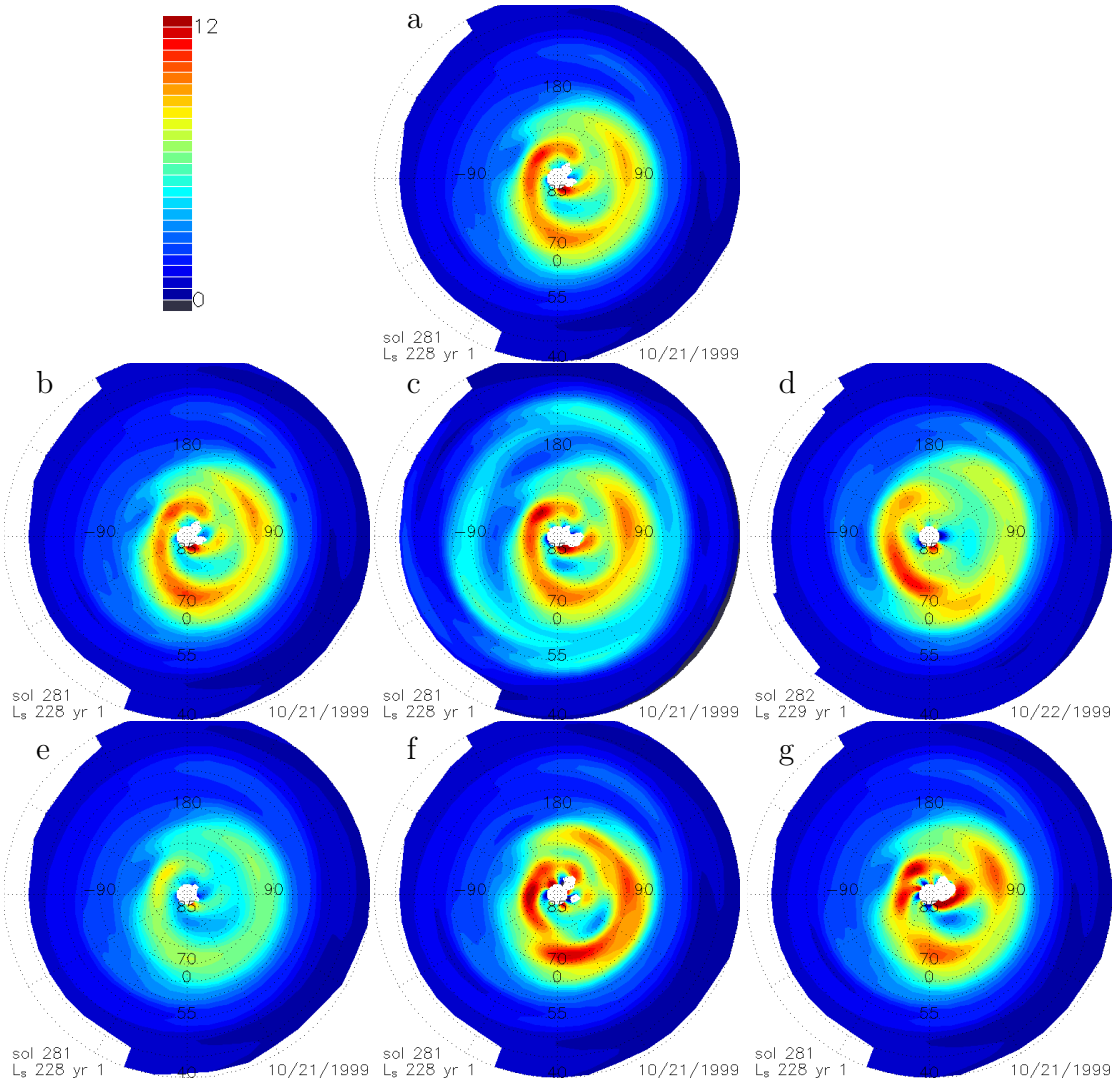

Figure 2.9: *PV* on the $\theta = 240$K surface, for the same sol and using the same map projection as in Fig. 2.3. The units on these and all subsequent *PV* plots are $10^{-4}\,\mathrm{K\,kg^{-1}\,m^2\,s^{-1}}$. (a) gives the preferred solution for potential vorticity, while (b) through (g) represent tests of the sensitivity of the solution to various assumptions and possible errors, as described in the text. (b) gives *PV* after the winds have been perturbed by twice the wind-solution residuals; (c) gives *PV* for the alternative 9.50 mb lower boundary; (d) gives the *PV* solution derived from nighttime rather than daytime TES observations; (e) gives *PV* after the temperature data set has been perturbed by a global systematic error of 1.2 Watt $\mathrm{cm^{-2}\ sr^{-1}\ (cm^{-1})^{-1}}$ as described in the text; (f) gives *PV* after the data set has been perturbed by exaggerated random errors (2.0 Watt $\mathrm{cm^{-2}\ sr^{-1}\ (cm^{-1})^{-1}}$ RMS) as described in the text; and (g) gives *PV* after each orbit of the temperature data set has been perturbed by 0.5 Watt $\mathrm{cm^{-2}\ sr^{-1}\ (cm^{-1})^{-1}}$ RMS random errors as described in the text. Note that, for reasons discussed in the text we show in (d) data from one time-step later than the other maps in this figure.

2.2.8 Modified Ertel Potential Vorticity

When plotting *PV* in cross section, we use the *modified* Ertel potential vorticity suggested by Lait [1994] in order to remove the inherently strong vertical variation in *PV*. Modified ertel potential vorticity, *PV′*, has the same conservation properties as *PV*, and is defined as [Lait, 1994]

$$PV' = PV\left(\frac{\theta}{\theta_{\text{ref}}}\right)^{-\left(1+\frac{c_p}{R}\right)} \tag{2.17}$$

2.2.9 Trajectories

We plot trajectories by numerically integrating our derived horizontal wind field on a θ-coordinate surface. The trajectory integration is accomplished by repeating the following procedure:

1. For a given latitude ϕ_n and longitude λ_n, determine the u_n and v_n wind components by bilinear interpolation from the gridded wind components. Also, calculate, again using bilinear interpolation, other desired quantities such as *PV* or pressure at the given latitude and longitude.

2. Move a certain distance to find the next latitude - longitude point:

$$\begin{aligned} \phi_{n+1} &= \phi_n + v_n \cdot timeStep_n \\ \lambda_{n+1} &= \lambda_n + u_n \cdot timeStep_n \end{aligned} \tag{2.18}$$

   where $timeStep_n$ is the largest value which satisfies:

$$v_n \cdot timeStep_n < 0.25^\circ\,; \quad u_n \cdot timeStep_n < 2.5^\circ\,; \quad timeStep_n < \frac{1}{24}\text{sol}\,. \tag{2.19}$$

The trajectories that result from this process are intended as a conceptual tool — a qualitative representation of the paths that particles might follow. It is difficult

to assess the quantitative accuracy of the trajectories, because we do not know what processes might be occurring on time scales smaller than our one-sol time step, and because of the large residuals in our v-component wind solution.

When plotting a time series of maps that include trajectories, we chose an arbitrary starting point and starting time for each trajectory, and then perform the above integration procedure both forwards and backwards in time to determine the particle path for the entire time series. Occasionally, the trajectory enters a large region of missing data, in which case no further integration is possible and the "particle" is treated as being frozen in place. Thus, in our trajectory plots, an arrowhead representing a particle will sometimes end up frozen in place within one of the white regions that represents missing data.

2.2.10 Zonal-mean quantities

In order to plot zonal-mean quantities as a function of L_s, we sort our map grids into 1° wide L_s bins, and then, for each latitude grid position, average all valid data points within the bin. After the data for a particular quantity have been binned in this manner, the results for each latitude are smoothed in the time (L_s) dimension using a 7° boxcar window.

2.2.11 CO_2 saturation temperature

We calculate the CO_2 saturation temperature using:

$$T_{\mathrm{sat}} = \frac{3182.48}{23.3493 - \ln\left(0.9532 \times p_{\mathrm{atm}}\right)} \tag{2.20}$$

where p_{atm} is the atmospheric pressure given in millibars, and T in Kelvin. The coefficients 23.3493 and 3182.48 are taken from James et al. [1992]. 0.9523 is the

mole-fraction of CO_2 in the martian atmosphere [Owen, 1992].

2.2.12 Aerosol optical depth

Column-integrated dust and water-ice aerosol optical depths are retrieved from TES data as described by Smith et al. [2000] and by Pearl et al. [2001], respectively. These optical depths are gridded in the same manner as the temperature data, and also binned into zonal means in the same manner is the temperature data, except that, of course, the column-integrated optical depth grids have no vertical dimension.

2.3 Results

Figures 2.10– 2.20 represent a survey of Mars polar vortex characteristics over a period of two-and-a-half martian years. After this survey, we illustrate polar vortex dynamics with a set of sample time series in Figs. 2.21 – 2.31. In this paper, we use the "Mars Year" (MY) designation that originated with Clancy et al. [2000]. This system arbitrarily defines MY 1 to begin with L_s 0° on April 11, 1955. MGS mapping began in March 1999, MY 24 $L_s \sim 100°$. Our data cover MY 24, 25, 26, and 27, which began in July 1998, May 2000, April 2002, and March 2004, respectively. In order to describe specific sols, we arbitrarily define sol 0.0 to be at spacecraft clock time ("SCLK") 600,000,000. Sol 1.0 is 88775 seconds (one Martian solar day) later, and so on. Since SCLK counts seconds from January 1, 1980, sol 0.0 occurs on January 5, 1999. As previously discussed, all maps are calculated for a specific instant in time. Since our time sampling interval is 1 sol, all maps have an integer-valued sol number.

2.3.1 Zonal-mean behavior

The gross features of the Mars polar vortex are evident from plots (Figs. 2.10 and 2.11) of zonal-mean quantities versus season (L_s). Most notably, and in striking contrast to the terrestrial case, the martian *PV* field is almost always annular in shape. *PV* reaches a maximum between 60° and 80° latitude (north or south), and then decreases towards the pole. On Earth, *PV* generally increases monotonically towards the pole (Hartmann [1983] discusses the implications of the exceptions to this rule), and even when the *PV* field is highly distorted, it never achieves an annular shape. The Mars southern hemisphere polar vortex in fact occasionally achieves multiple maxima in zonal-mean *PV*, for example near 30° L_s in MY 26 and 27. Double-ring structures actually occur sporadically throughout the southern hemisphere winter in all three years, as well as in late winter in the northern hemisphere, but most of these events are too transient to show up in the zonal mean.

Second, the seasonal evolution of the area encompassed by the northern hemisphere polar vortex (considering the zonal-mean *PV* maximum to be the boundary) is distinctively asymmetric, in contrast to the terrestrial polar vortices, whose area steadily increases towards midwinter, and then shows a symmetric decrease into spring [Waugh and Randel, 1999]. On Mars, the northern vortex area is at a minimum in early winter near L_s 230°, and then begins to increase. It continues to increase past the solstice until about 310° L_s, after which the three winter periods show no particularly consistent behavior. Note that, since Mars's radiative time constant is short [$\sim$ 2 sols; Zurek et al., 1992]), we consider the solstices to be effectively the middle of the winter period on Mars.

In contrast to the Mars southern hemisphere polar vortex, the northern polar

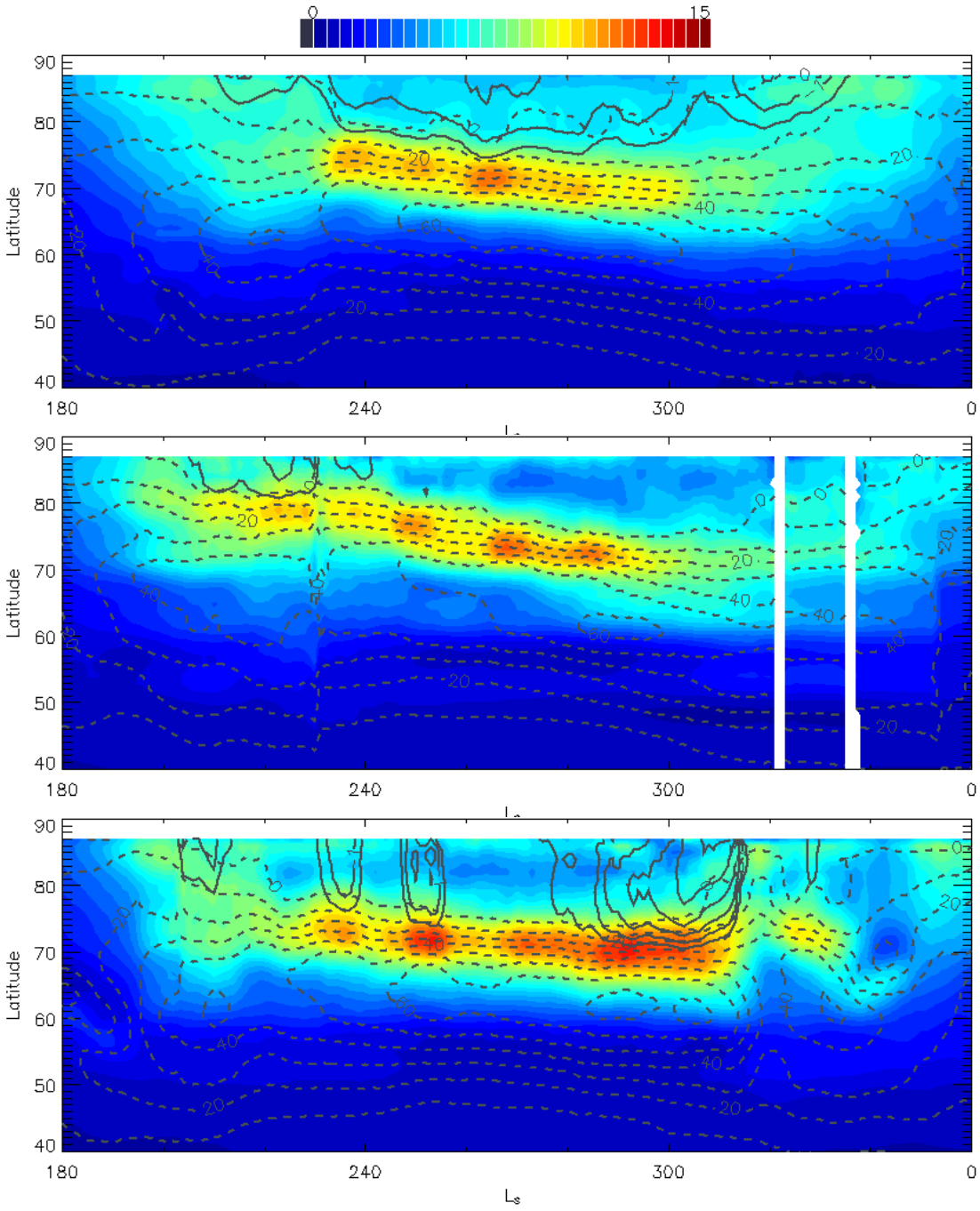

Figure 2.10: Zonal mean *PV* (colors) overlain with zonal mean temperature (solid lines) and zonal winds ($\mathrm{m\,s^{-1}}$, dashed lines) for northern hemisphere winter in (top to bottom) MY 24, 25, and 26. All plotted data is for the $\theta = 240$K surface. Temperatures are shown as the difference between the measured temperature and the CO_2 saturation point — we plot contours only where CO_2 is supersaturated ($T - T_{\mathrm{sat}} \leq 0$), in one K intervals starting from zero. *PV* has units of $10^{-4}\,\mathrm{K\,kg^{-1}\,m^2\,s^{-1}}$.

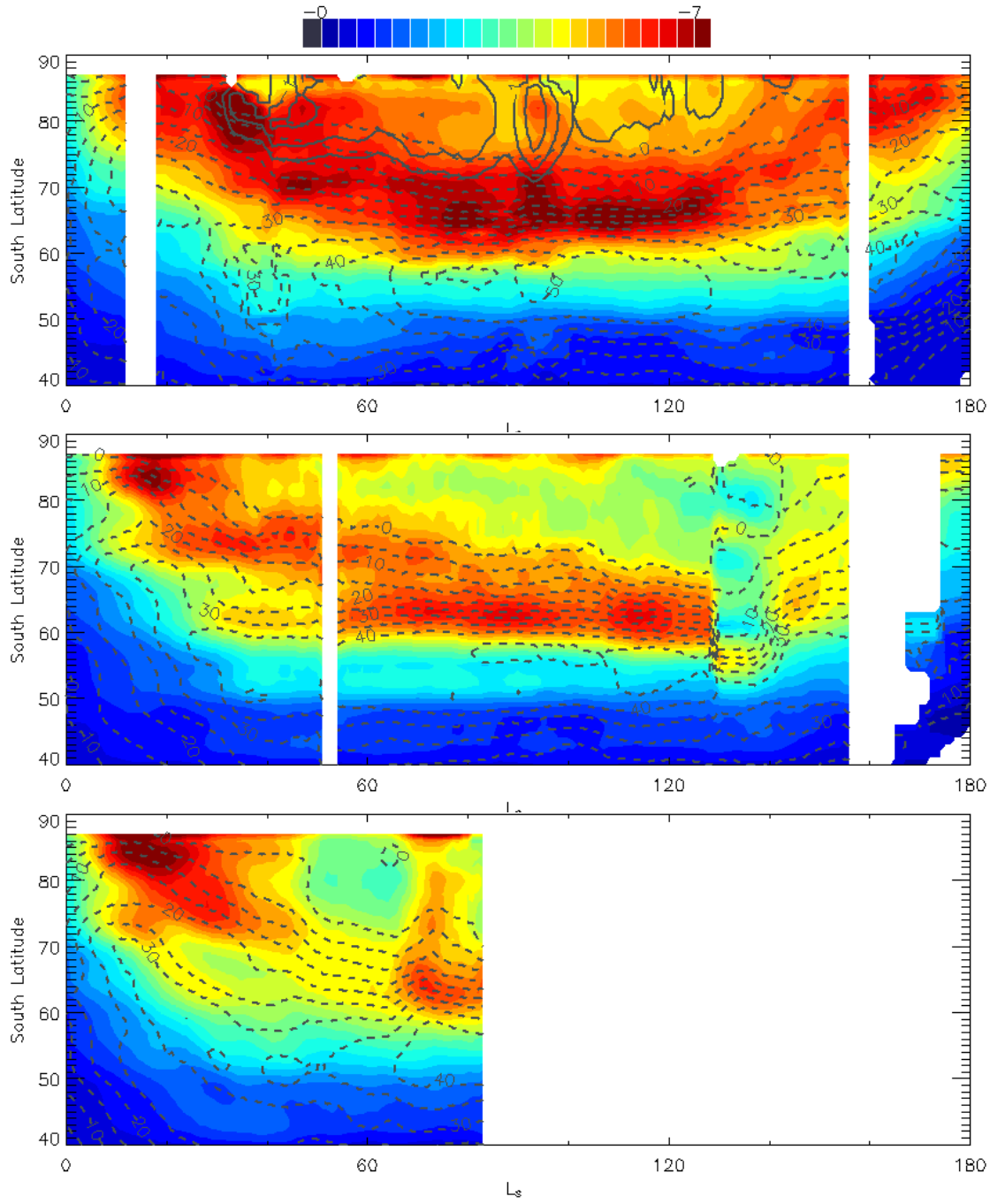

Figure 2.11: Zonal-mean *PV* (colors) overlain with zonal mean temperature (solid lines) and zonal wind ($\mathrm{m\,s^{-1}}$, dashed lines) for southern hemisphere winter in (top to bottom) MY 25, 26, and 27. All plotted data is for the $\theta = 240$K surface. Temperatures are shown as the difference between the measured temperature and the CO_2 saturation point — we plot contours only where CO_2 is supersaturated ($T - T_{\mathrm{sat}} \leq 0$, in one K intervals starting from zero. *PV* has units of $10^{-4}\,\mathrm{K\,kg^{-1}\,m^2\,s^{-1}}$.

vortex is generally more intense, in the sense that the *PV* field is stronger. The high-*PV* region in the north is more compact, and the maximum *PV* is closer to the pole. The position of the *PV* maximum at winter solstice seems to be consistent — near 64° latitude in the south, and near 72° latitude in the north. The previously described pattern in the north pole vortex area was also similar in each year. Otherwise, the *PV* pattern is quite variable from year to year in each hemisphere.

In the south, for MY 25, *PV* evolution was roughly symmetric about the solstice (L_s 90°), but MY 26 had a completely different, and asymmetric, pattern. That year exhibited the previously mentioned double-ring pattern, a fairly constant boundary latitude from L_s 60° to L_s 120°, and then a sudden drop in *PV* intensity near L_s 130°. This *PV*-field breakdown was followed by redevelopment of a weaker polar vortex at L_s 140°. MY 27 in the south, like MY 26, shows a period of multiple *PV* maxima near L_s 30°, but unlike the other years, the *PV*-field seems to grow notably weaker during L_s 50°–65° before strengthening again.

In the north, the most prominent inter-annual variations are the size and *PV*-field intensity of the polar vortex in early winter, and the intensity of the *PV*-field in late winter. The MY 25 vortex boundary at L_s 230° was near 79° north latitude, versus 74° in the other years, and was followed by a brief apparent *PV*-field breakdown. The compressed MY 25 vortex was also more intense than the other two years until about L_s 230°. The MY 26 *PV* maximum at L_s 300° was about 40% greater than in the other years. The excess-intensity period in MY 26 preceded an extended *PV*-field breakdown which began near L_s 310°. This breakdown ended shortly after L_s 320° with the re-emergence of an annular *PV* field that was also much more intense than observed in the same season of the other years.

Northern winter *PV* variability shows some correlation with the low-latitude

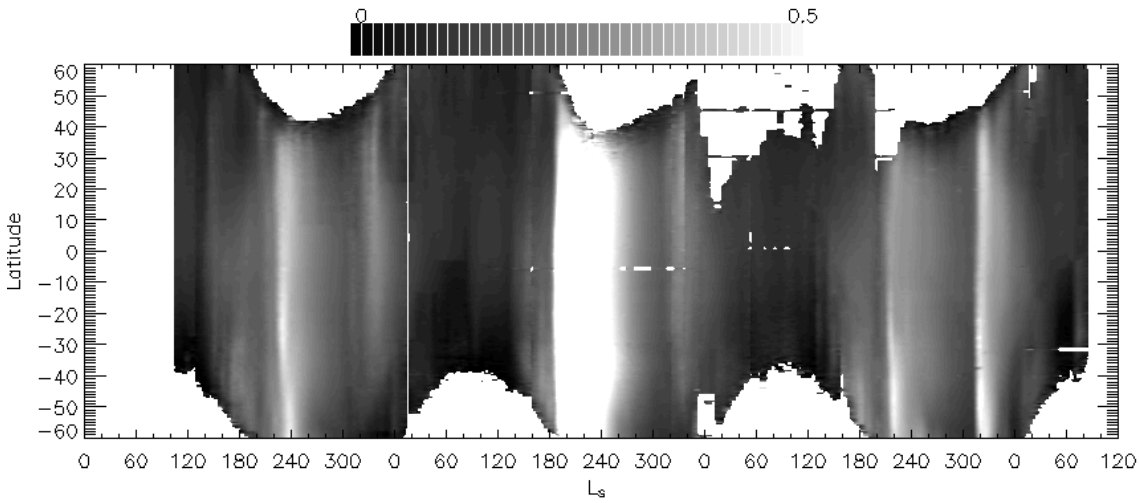

Figure 2.12: Zonal-mean dust opacity (as in Smith et al. [2000]) for the martian years shown in the previous figures.

dust opacity ($\tau_{\rm dust}$, Fig. 2.12). The period of minimum vortex area near L_s 230° is associated with a high-dust period in all cases; the more compact early winter vortex of MY 25 occurred during the planet-encircling dust-storm of 2001; and the MY 26 breakdown coincides with yet another high-dust period. However, some of the *PV* variability is apparently not connected to dust — all three years had similar dust levels, but quite different *PV* levels, around L_s 300°. In the southern hemisphere, dust doesn't seem to be a factor in *PV* variability. Dust opacity is roughly the same and very low in this data set's two southern winters.

In addition to *PV*, Figs. 2.10 and 2.11 show contours (dashed lines) of zonally averaged eastward wind, and also highlight (with solid-line contours) regions where the zonal mean retrieved temperatures reach the CO_2 saturation point ($T - T_{\rm sat} \leq 0$). The wind contours provide information about the strength and location of the polar vortex that is similar to that provided by *PV*, so their value is primarily as an aide to visualizing the relationship between the polar jet and the *PV* field. Thus, these figures illustrate that *PV* is closely associated with the wind shear, and

demonstrate the mathematical requirement that the high-*PV* region must lie on the poleward flank of the jet. Not surprisingly, the maximum speed in the jet is highly correlated with the shear on its poleward flank and thus with *PV*.

The $T - T_{\rm sat}$ contours show that persistent zones of supersaturation occur in the MY 24 and MY 26 northern winters, as well as in the MY 25 southern winter. Given the small departure from the saturation point, and the possibility of periods of systematic errors in the TES temperatures, the apparent supersaturation is not a robust result, but it does suggest where and when atmospheric condensation is most likely to be occurring. Note also that the the polar vortex high *PV* zone appears to confine these cold regions, especially in northern winter of MY26.

In order to illustrate the more general connections between temperature and the dynamical state of the vortex described by the *PV* field, we show zonally averaged temperatures for several polar latitudes in Figs. 2.13 and 2.14. The low-latitude temperature variations in northern winter are mainly, as expected, a reflection of dust loading. In contrast, at 80° north, which is almost always within the vortex's high-*PV* ring, the zonal temperatures usually show no response to dust. The one exception is a major warming in conjunction with the MY 26 L_s 300° *PV*-field breakdown. The brief breakdown in MY 25 is not associated with any polar warming.

The southern *PV*-field breakdown in MY 26 is also associated with a significant polar warming, relative both to temperatures immediately preceding the event, and to the same season in the preceding year. Polar (80° south) temperatures are at least 5 K warmer than the previous year for most of the winter, while lower (closer to the equator) latitude temperatures are very similar to those in the preceding year. The warming event in southern MY 26 differs from the northern MY 26 event

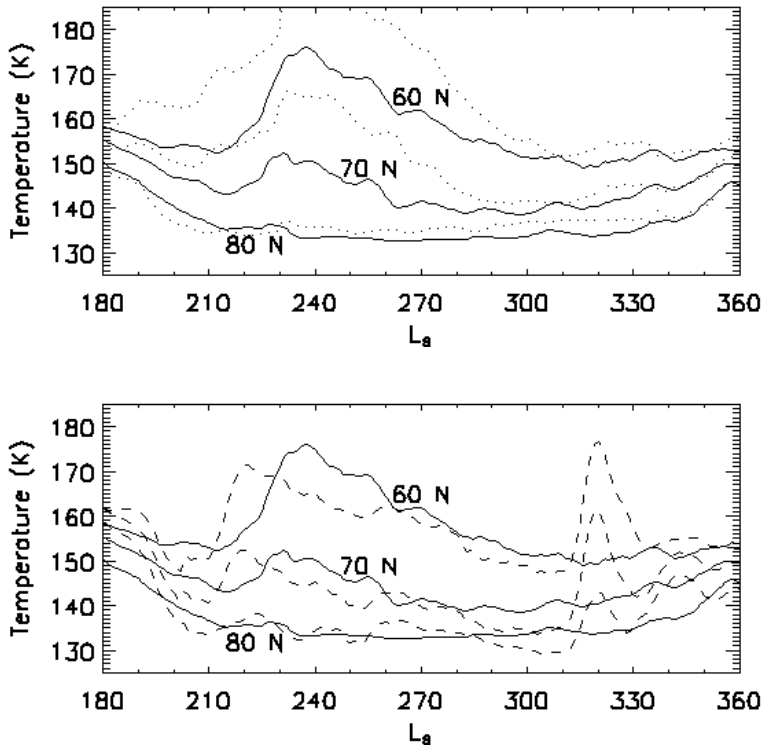

Figure 2.13: Temperature at three different northern latitudes on the 0.83 mb surface for MY 24, solid lines; MY 25, dotted lines; and MY 26, dashed lines. The plot on the top shows year 24 and year 25. The plot on the bottom shows MY 24 and MY 26.

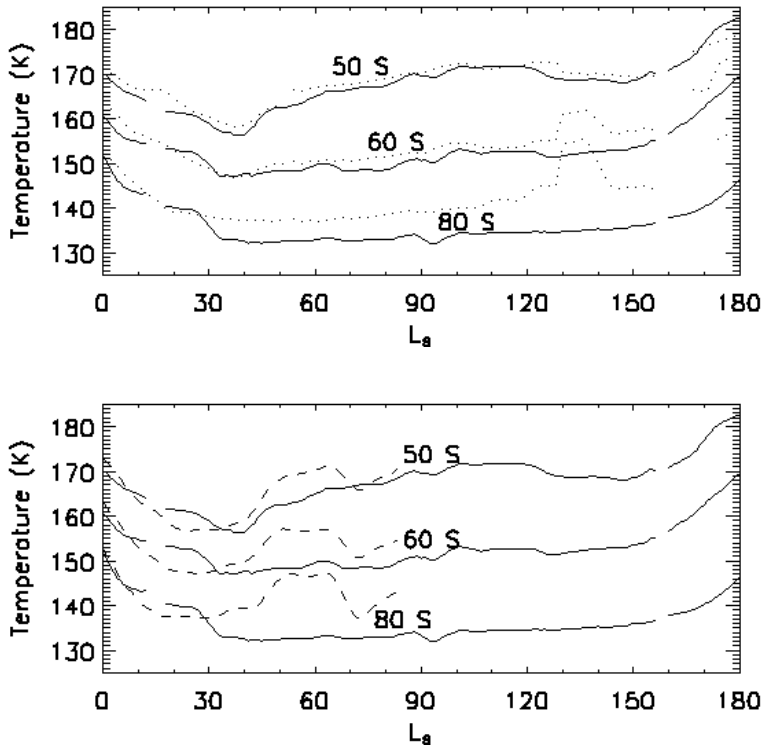

Figure 2.14: Temperature at three southern different latitudes on the 0.83 mb surface for MY 25, solid lines; MY 26, dotted lines; and MY 27, dashed lines. The plot on the top shows MY 25 and MY 26. The plot on the bottom shows MY 25 and MY 27.

in that the warming is confined to the polar region, with temperatures actually decreasing slightly at 50° south latitude.

For the southern MY 26 warming, the low-latitude decrease in temperature rules out the possibility that it is caused by a systematic error in TES radiances, because a systematic radiance error would produce a radiance offset in the same direction at all latitudes. A systematic error can also be confidently ruled out for the northern MY 26 warming, partly because of its correlation with dust opacity, but also because the temperature offset of a radiance error is inversely ($\sim T^{-4}$) related to the true temperature, and the magnitude of this northern warming is much greater at the warmer lower latitudes. However, the southern MY 27 warming *is* consistent with a radiance error. If we take the MY 25 temperatures to be the "true" temperatures, we find the temperature offsets have the signature T^{-4} scaling. Since the reality of this event is in doubt, we will not analyze it further in this paper.

2.3.2 Spatial pattern

Figs. 2.15 – 2.20 present spatially resolved information about the polar vortex structure for selected sols in early (top of figure), mid-, and late (bottom of figure) winter, showing both modified potential vorticity, PV', and temperatures where T is below the saturation point. The cross sections in these figures also show lines representing surfaces of constant entropy ($\theta =$ constant), and the maps give PV' and temperatures below the saturation point for one of these surfaces. Note that, by our definition (Eq. 2.17), PV' is just a constant multiple of PV on any constant-θ surface. Using PV' merely makes it easier to compare PV structure on different θ surfaces. For the map views, we have chosen the theta surface on which PV is *equal*

to PV'.

These figures show that the PV field is in fact commonly annular as inferred from the zonal means, and that this annular character persists throughout the vertical range of our data set. In the northern hemisphere midwinter period, the annulus is remarkably organized and symmetric, although even in this period it is often distorted by wave activity as seen in Fig. 2.17d–f and the time series of Fig. 2.22. The early winter period in the north also tends to be relatively organized. This organized behavior is a key contrast with the southern polar vortex. There, the zonal mean annulus is made up of arcuate, sometimes spiral-like, high-PV clumps, and, as the cross sections shows, there may be multiple maxima along any given meridian. As an extreme example, Fig. 2.19c — early southern winter during the previously mentioned zonal-mean double ring — shows four local maxima on most θ-surfaces. But generally disorganized and asymmetric behavior with multiple maxima persists throughout southern winter, even though a zonal-mean double ring is not observed, and is also the rule in late northern winter. Ironically, the one exception in the maps we present (Fig. 2.18a-c) occurs at the same L_s as the double ring, but in the preceding year. At this time the PV increase was nearly monotonic towards the winter pole.

Figures 2.15 – 2.20 also illustrate the zones of symmetric instability ($PV < 0$ in the northern hemisphere, $PV > 0$ in the southern hemisphere) noted by Banfield et al. [2004]. These zones are generally larger, and reach down to lower pressure levels, in the northern hemisphere, especially near solstice. In the southern hemisphere, there appear to be symmetrically unstable regions very near the pole, although these could conceivably be instances of the wind solution breaking down near the edge of the domain.

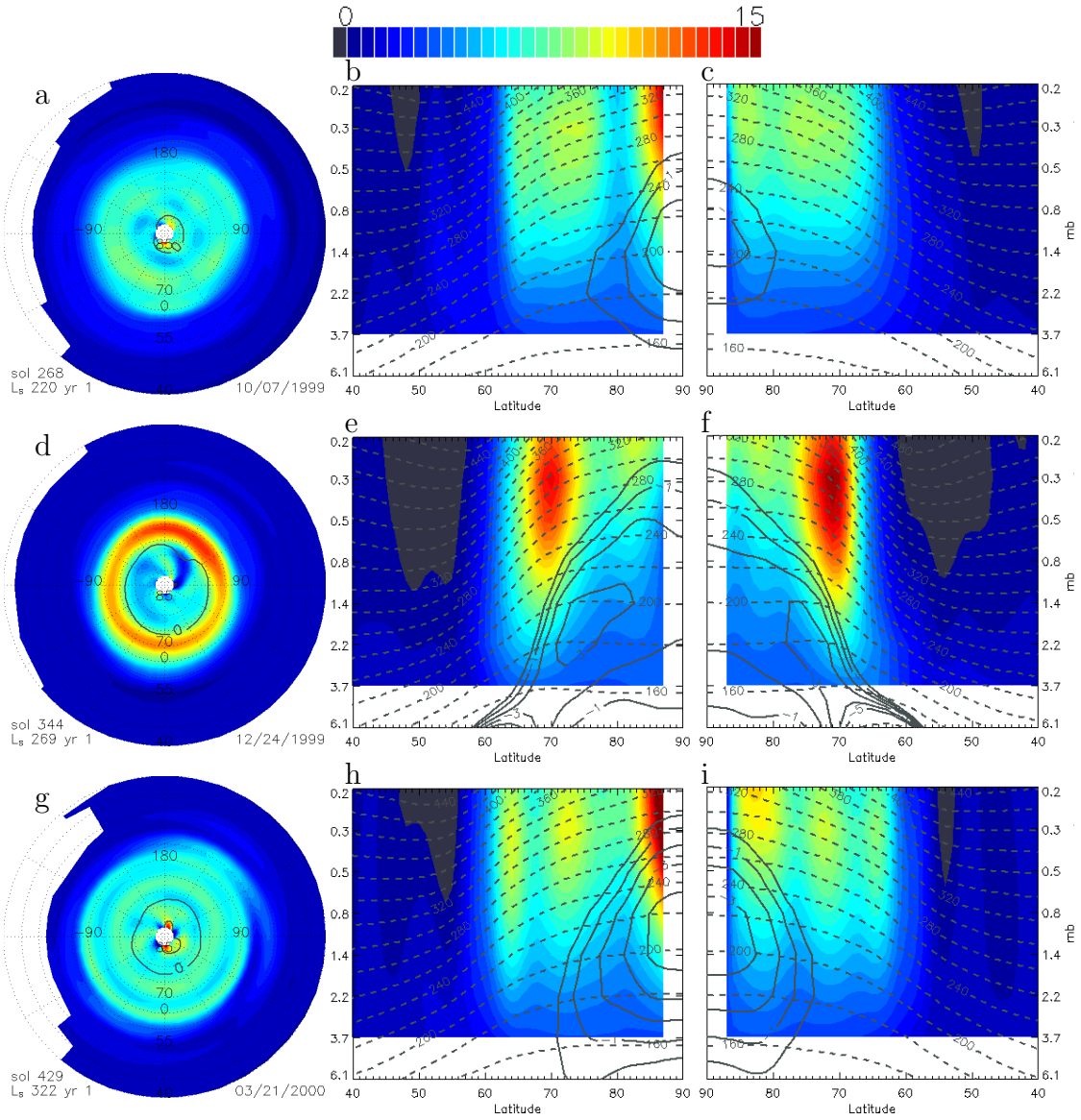

Figure 2.15: Maps ($\theta = 240$K surface), (a), (d), and (g) and cross sections for selected sols in (top to bottom) early, mid-, and late northern hemisphere winter — MY 24. Left cross-sections, (b), (e), and (h), show quantities averaged over 0°–30° east, and right cross sections, (c), (f), and (i), show quantities averaged over 180°–210° east. The colors show modified potential vorticity, PV' ($10^{-4}\,\mathrm{K\,kg^{-1}\,m^2\,s^{-1}}$). Dark grey colors indicates regions of symmetric instability, $PV < 0$ in the northern hemisphere, $PV > 0$ in the southern hemisphere. The map projection is as established in Fig. 2.3. Solid lines show temperatures where $T - T_{\mathrm{sat}} \leq 0$. Dashed lines on the cross sections show potential temperature, θ.

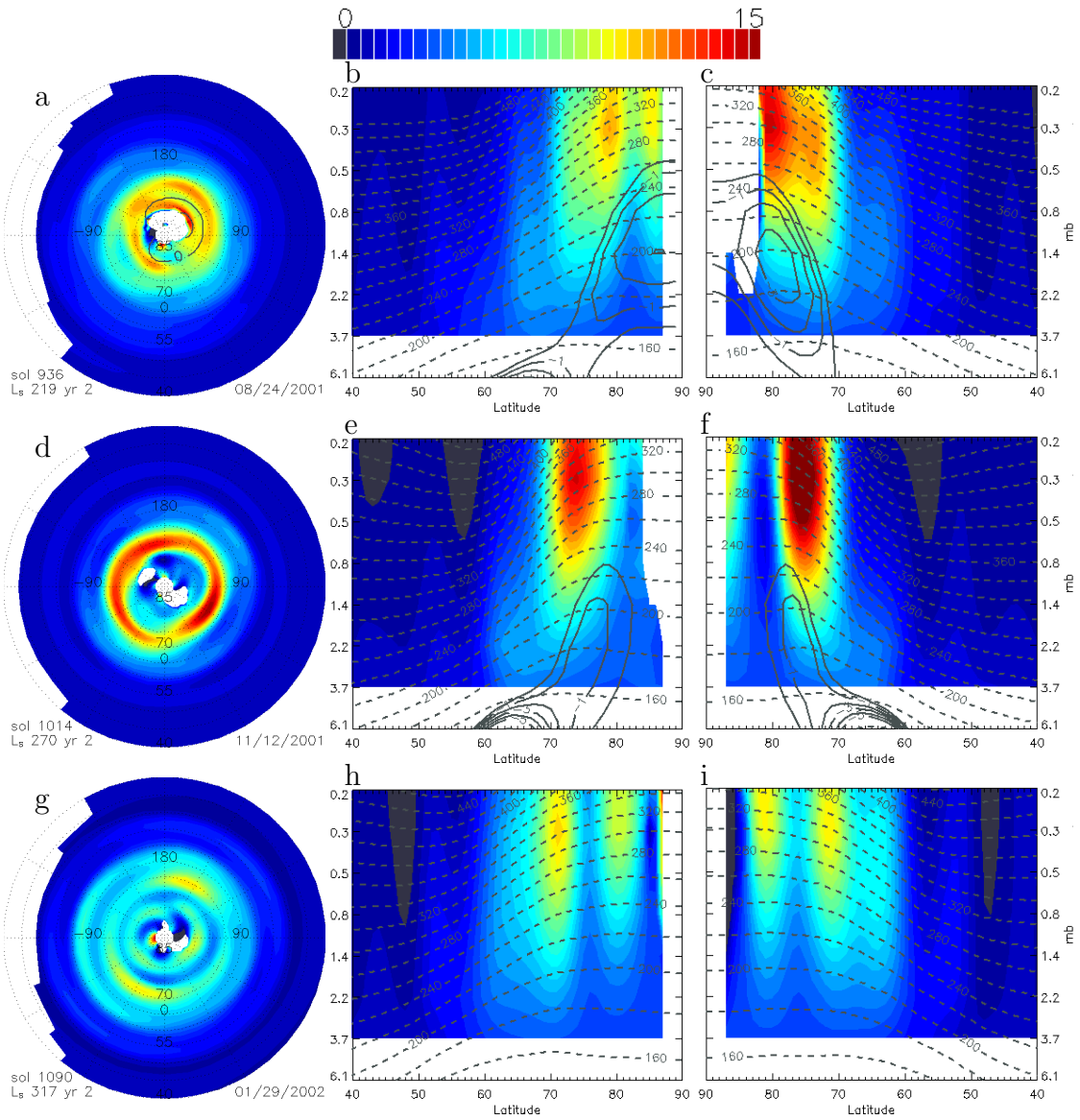

Figure 2.16: Same as Fig. 2.15, but for MY 25.

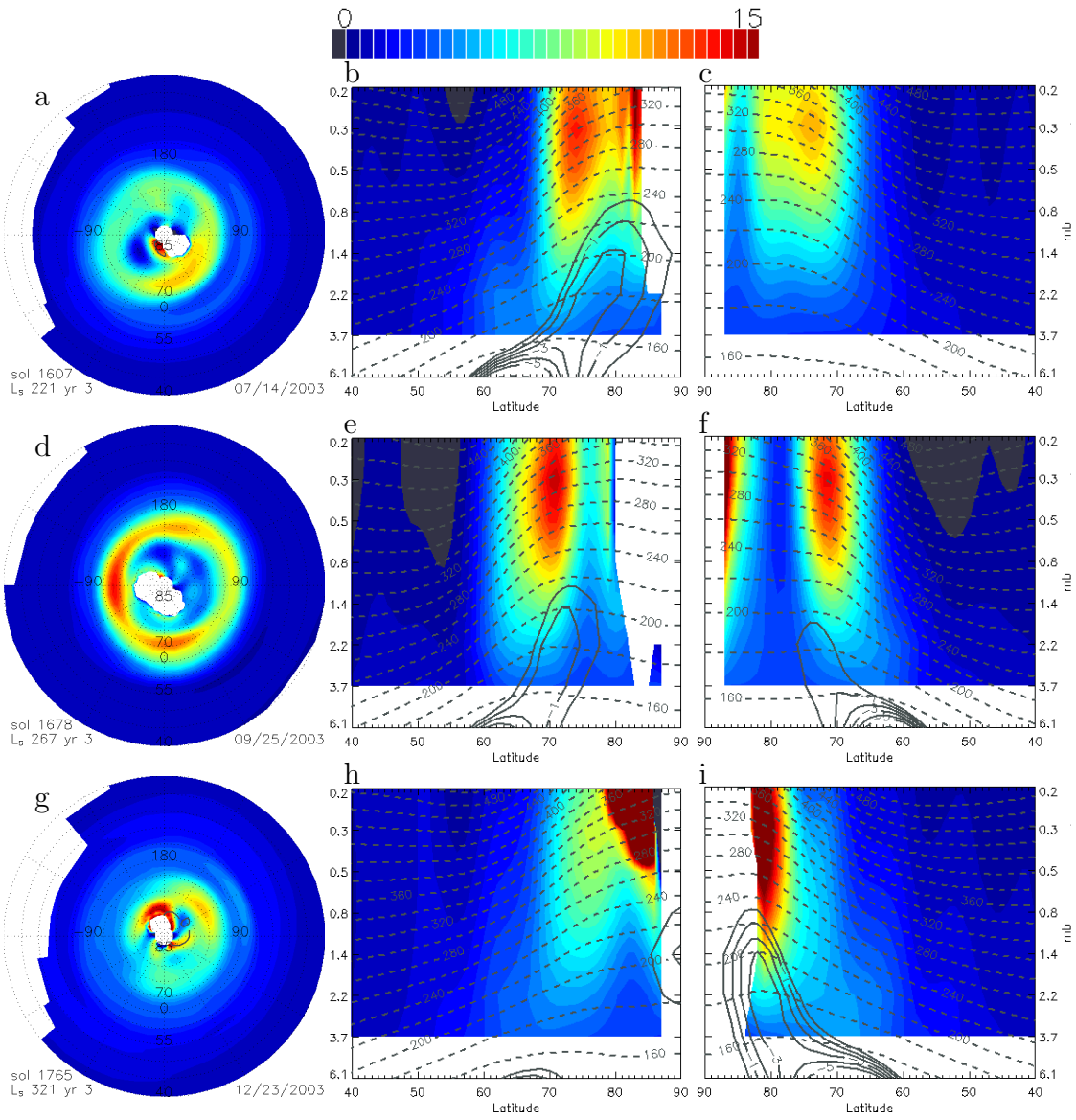

Figure 2.17: Same as Fig. 2.15, but for MY 26.

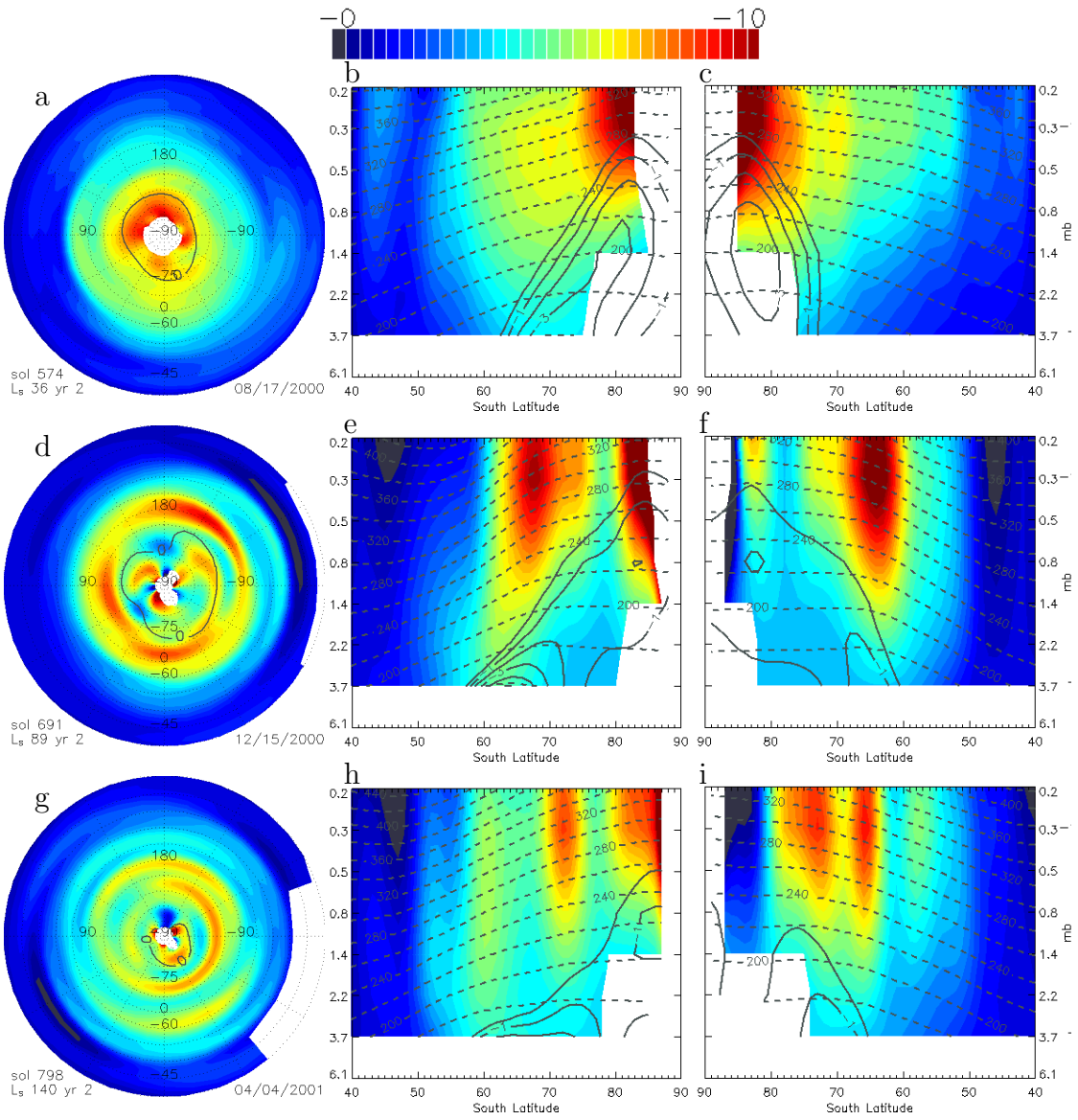

Figure 2.18: Same as Fig. 2.15, but for the south pole, southern winter, MY 25.

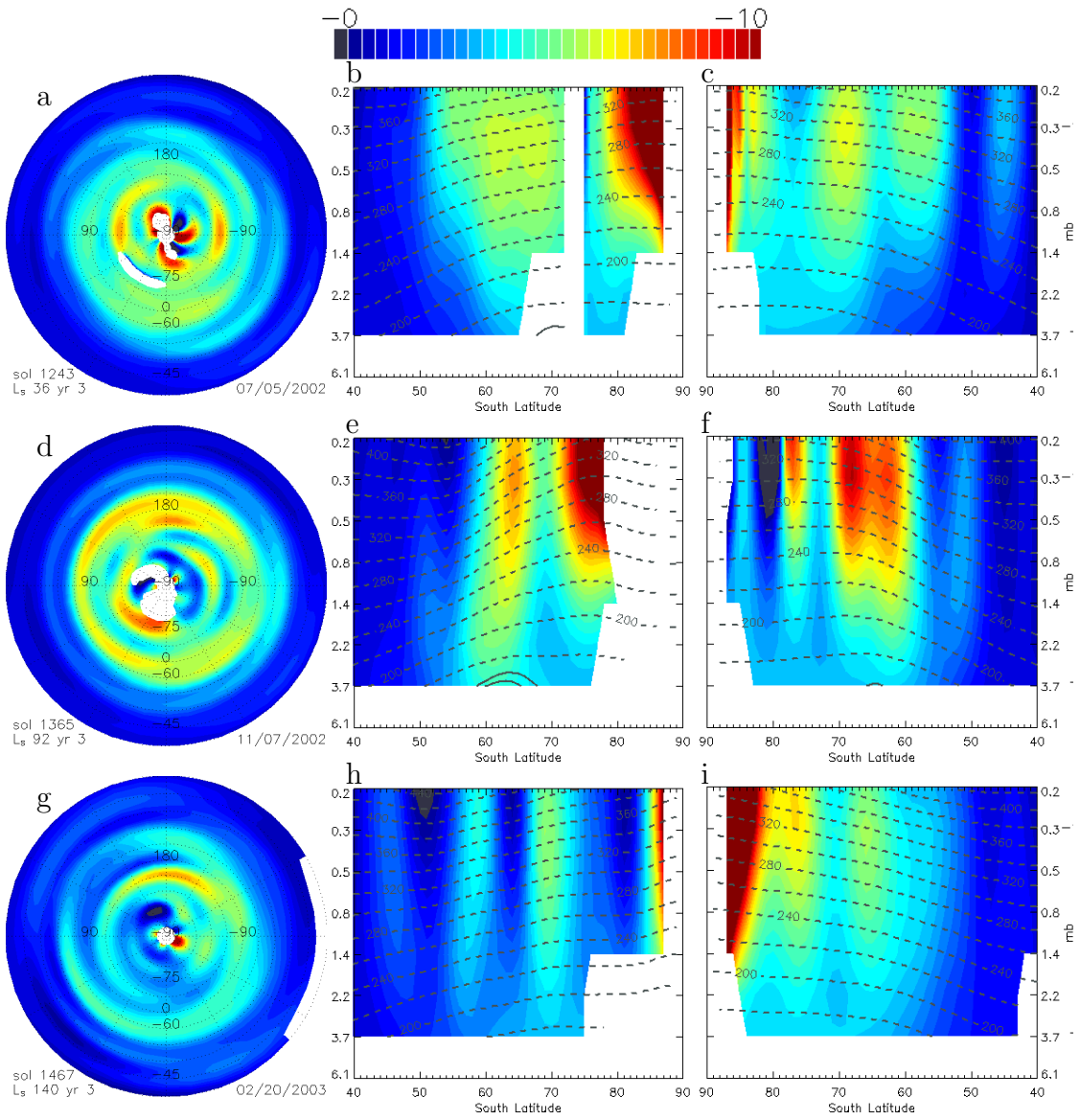

Figure 2.19: Same as Fig. 2.15, but for the south pole, southern winter, MY 26.

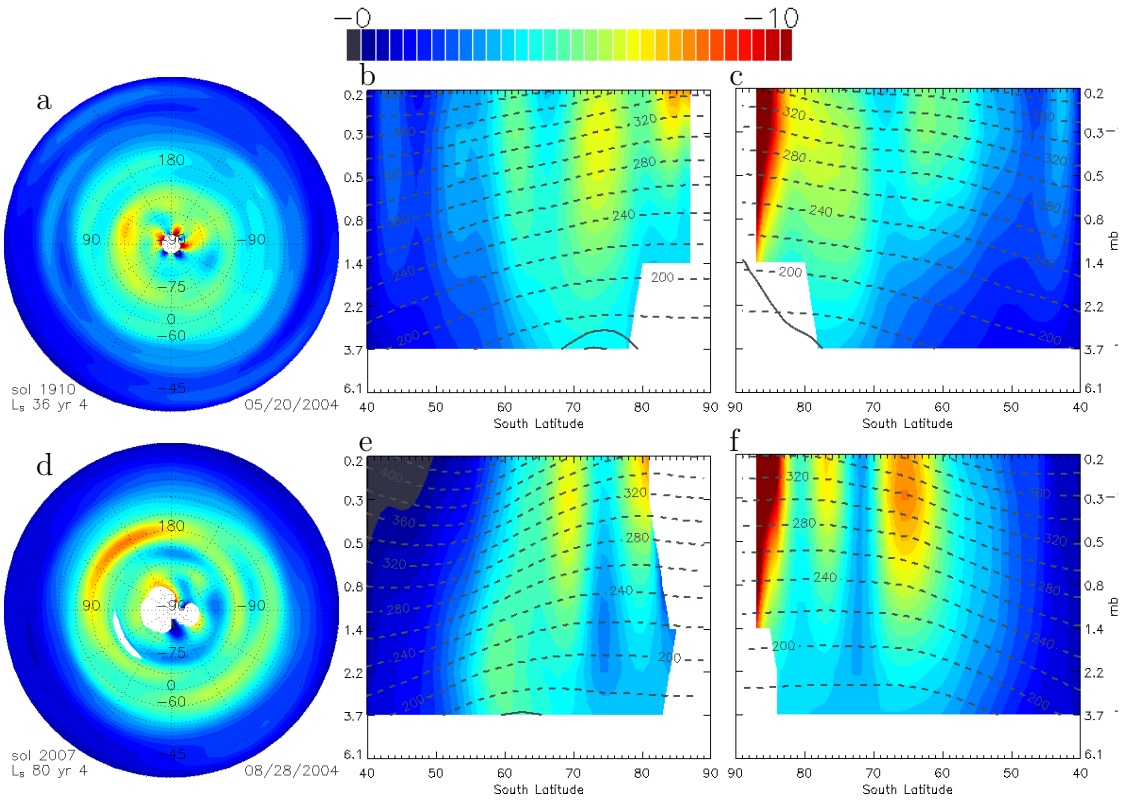

Figure 2.20: Same as Fig. 2.15, but for the south pole, southern winter, MY 27.

Figures 2.17g–i and 2.19g–i give additional insight into the nature of the *PV*-field breakdown events that we noted in the zonal mean. Both figures show conditions just before the *PV*-field reforms. For the northern hemisphere MY 26 breakdown, the *PV* annulus has been displaced and shrunken to the point where it is difficult to recognize in the zonal mean, but in fact the *PV* field is more intense, especially at high altitudes, and is better organized than at similar L_s in previous years in the sense that it has only a single *PV* maximum on each side of the pole. By contrast, during the southern hemisphere breakdown, the *PV*-field is notably weaker than the previous year at all altitudes, and shows what look like major spiral-like intrusions into the polar region by low-*PV* air.

Finally, Figs. 2.15 – 2.20 provide more detail about the location of the apparently supersaturated regions, and their relationship to the *PV* field. $T - T_{\mathrm{sat}}$ values as low as -5 K are observed, which suggests that the retrieved temperatures are too low in these regions. However, anecdotal reports by Colaprete et al. [2003] of CO_2 supersaturation in MGS radio occultation temperature profiles suggest that supersaturation in the martian polar regions is a real phenomenon, and in any case it is well established that not all of the CO_2 condensation occurs at the surface [Colaprete and Toon, 2002].

Regardless of the true $T - T_{\mathrm{sat}}$, the observed values provide important information about which regions and seasons are most likely to be affected by CO_2 condensation. In the northern hemisphere, a particularly clear pattern of saturation emerges. The lowest $T - T_{\mathrm{sat}}$ values occur in a narrow region between 60° and 70° and confined below the 3.7 mb pressure level. In other words, at the bottom of our data domain, the low $T - T_{\mathrm{sat}}$ zone forms an annulus which is equatorward of the high-*PV* annulus. The other zone of usually low $T - T_{\mathrm{sat}}$ values

is just *inside* (poleward of) the high-*PV* annulus, typically between the 2.2 and 0.5 mb pressure levels. These two zones are connected by a region of less intense apparent saturation, so that overall the more saturated air forms a sort of dome which arcs over a less saturated or unsaturated zone near the surface at the winter pole. Figure 2.15e–f is the type example of this pattern, with all of the other figures showing variations on it. When the *PV* field is weaker or less organized, the bottom-of-domain annulus of low temperatures is weaker or absent. Similarly, when the *PV* field is asymmetric on a given sol, the $T - T_{\mathrm{sat}}$ field is also asymmetric, with the lowest bottom-of-domain$T - T_{\mathrm{sat}}$ levels occurring on the side of the pole with the highest *PV* levels.

2.3.3 Time Series

Figures 2.21 – 2.24 show times series in order to provide a sampling of the dynamical behaviors of the Mars polar vortex at different times and different levels of the atmosphere. In addition to *PV*, each time series frame shows the location and path of a collection of test particles. Although these trajectories are only approximate, given the lack of precision in the v-component wind estimates, they provide a rough idea of how a tracer such as *PV* should be advected by the winds. Also, since the paths plotted in each frame show one full sol worth of motion, the length of the paths provide a direct indication of the Rossby number (provided that the center of curvature of the path is not too far from the pole). For example, a particle which covers 60° of longitude in one sol along a circle centered at the pole would have a Rossby number of exactly $\frac{60}{360} = \frac{1}{6}$.

In the higher level ($\theta = 300$) northern hemisphere time series (Figs. 2.22 and 2.24), the fastest trajectories have Rossby numbers between 1.0 and 1.3, whereas

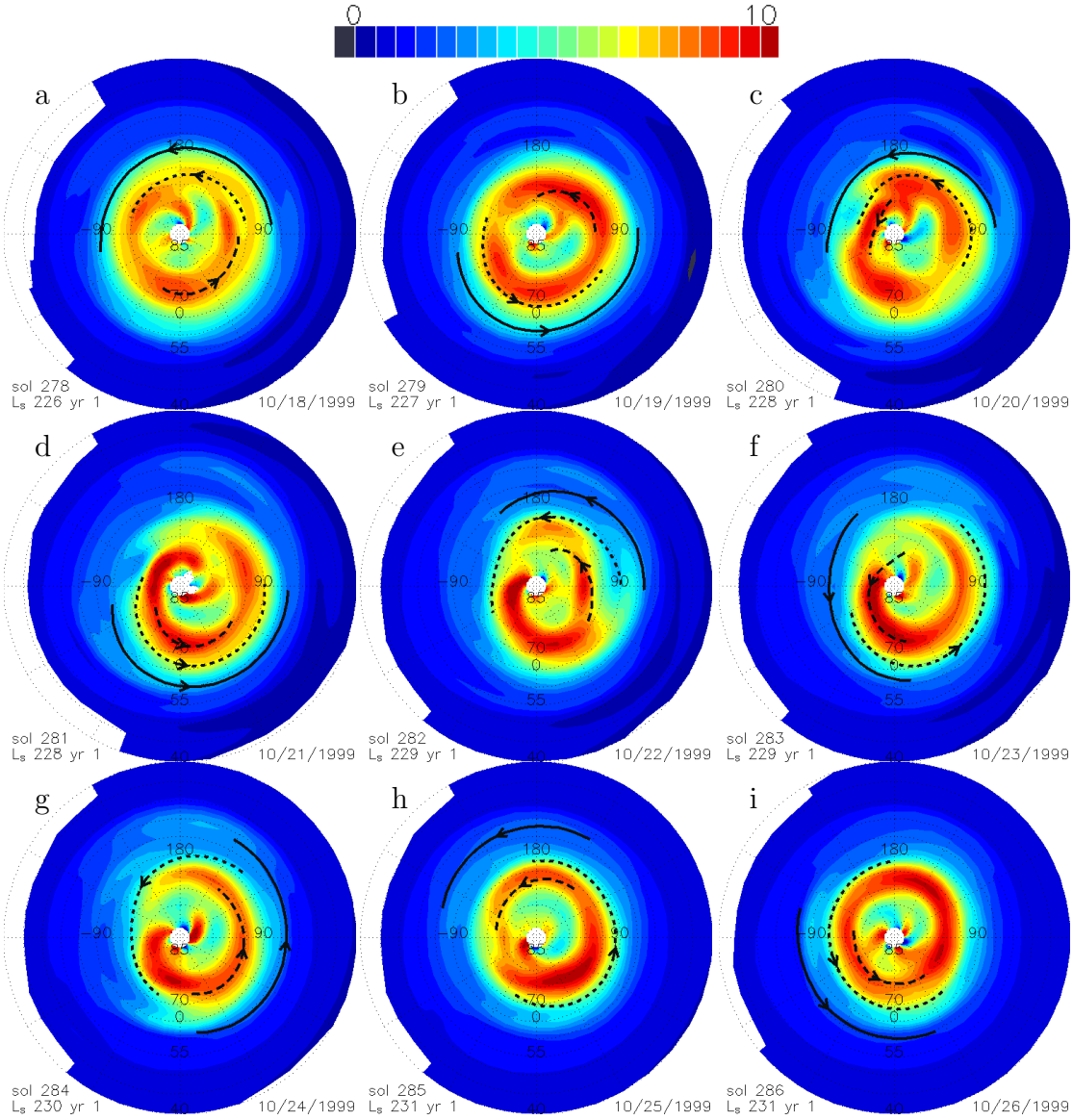

Figure 2.21: Time series of *PV* maps, sols 278 – 286 in the northern hemisphere, on the $\theta = 240$K surface, with map projection and units as previously described. Lines and arrows show calculated trajectories for several test particles, with a distinct line style (solid, dotted, dashed, etc.) for each test particle. The arrowhead marks the location of the test particle at the instant in time that corresponds to the *PV*-field shown in the map. Each line traces the path of the particle for one half sol before and one half sol after this instant.

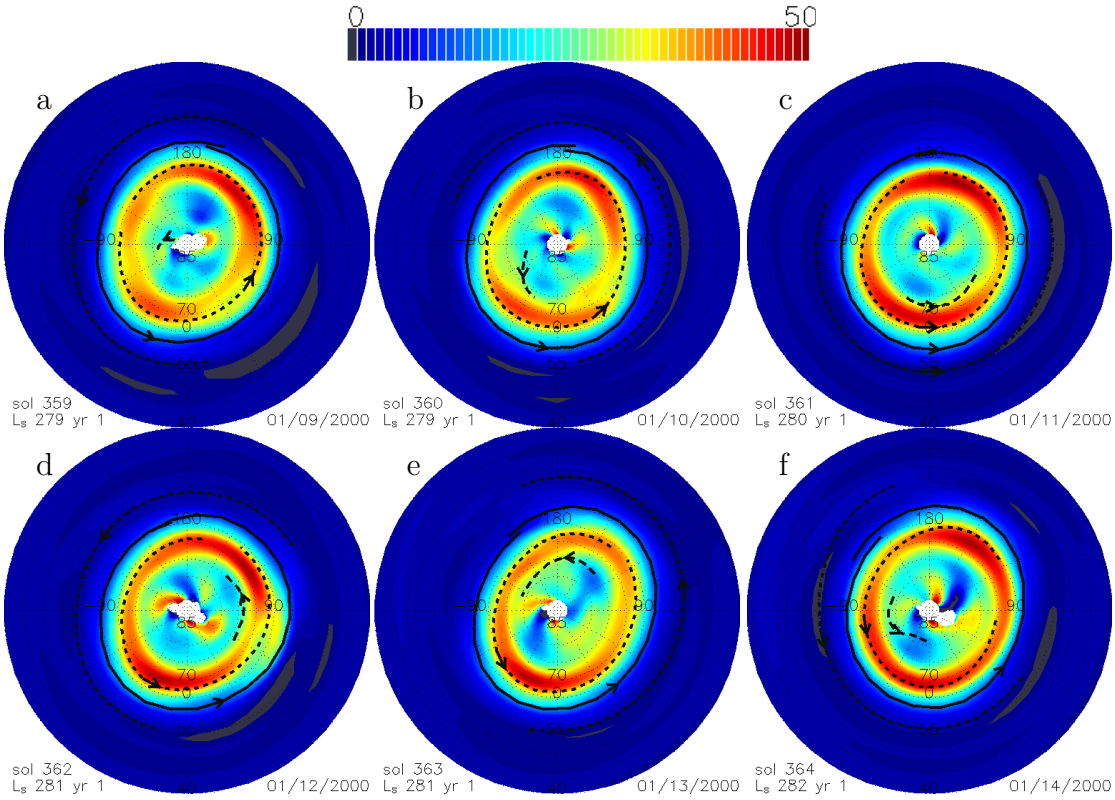

Figure 2.22: Time series of PV maps, sols 359 – 364 in the northern hemisphere, on the $\theta = 300$K surface. Otherwise the same as Fig. 2.21.

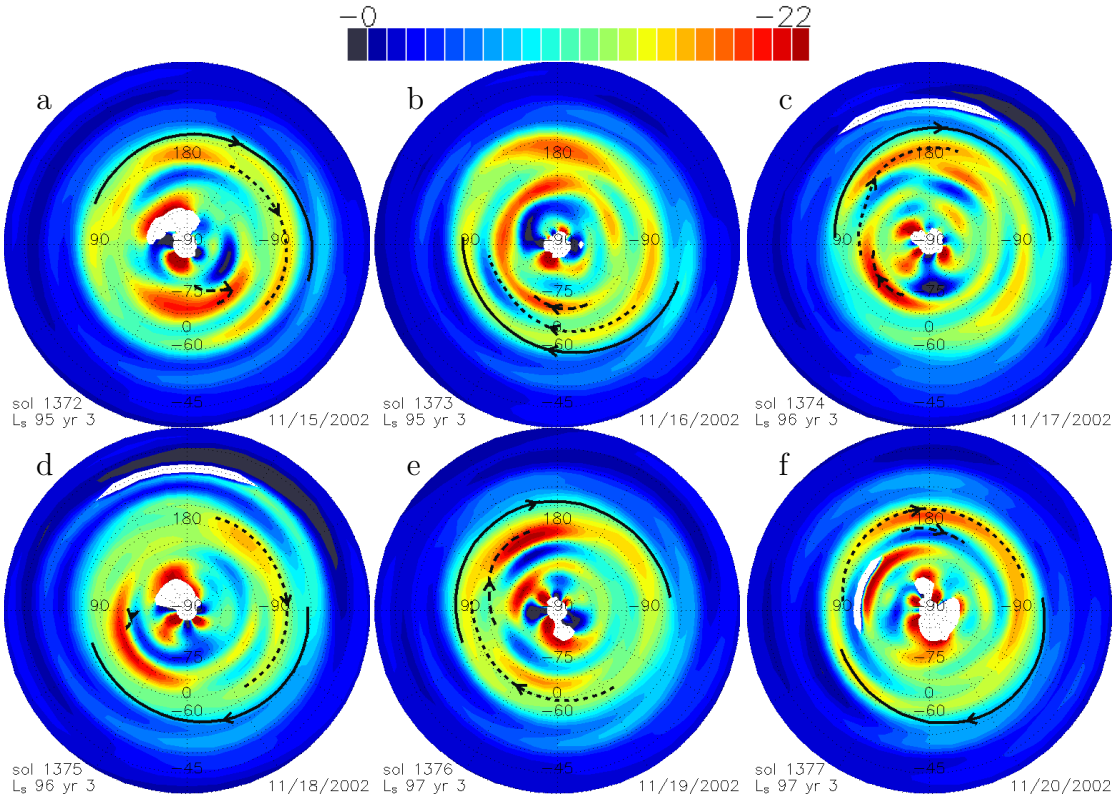

Figure 2.23: Time series of *PV* maps, sols 1372 – 1377 in the southern hemisphere, on the $\theta = 280$K surface. Otherwise the same as Fig. 2.21.

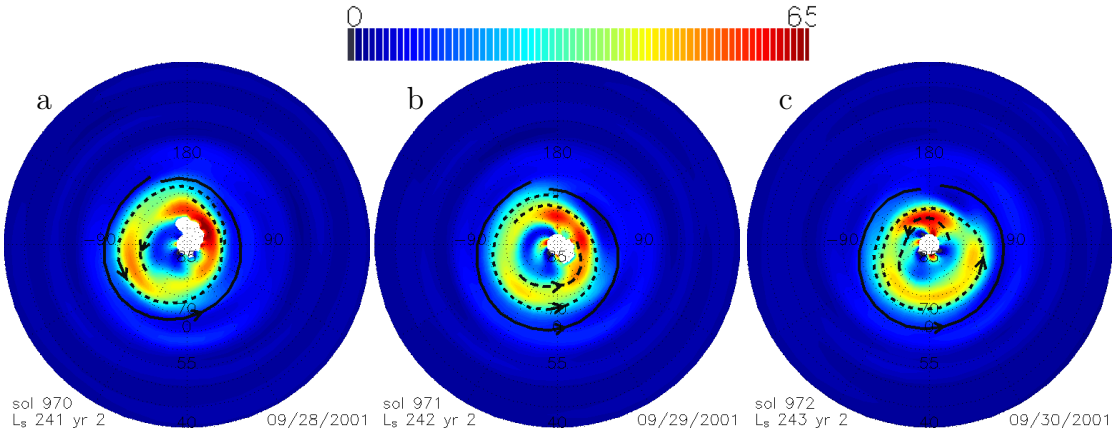

Figure 2.24: Time series of *PV* maps, sols 970 – 972 in the northern hemisphere, on the $\theta = 300$K surface. Otherwise the same as Fig. 2.21.

the other times series (a lower level in the north, Fig. 2.21, and a southern hemisphere series, Fig. 2.23, have Rossby numbers of 0.5 at the most. Rossby numbers this large obviously have important implications for analyses of polar vortex dynamics, but the high wind speeds, in and of themselves, also mean that a great deal of evolution takes places in the *PV* field in between the time steps of our analysis grid. Thus, it is very difficult to follow the path of any particular *PV* feature in order to, for example, assess whether low-latitude air is being mixed into the interior of the polar vortex.

The first time series we present (Fig. 2.21) shows an interesting spiral-like event from early northern winter in MY 24. It is "spiral-like" in the sense that the zone of maximum *PV* develops a spiral pattern which persists for about four sols before dissipating. At the onset of this event, both the *PV* maximum and the particle paths are displaced northward by about 10° near 270° East. Although the opening of the "spiral" is a gap in the ring of high-*PV* that separates polar air from equatorial air, the *PV* in that gap is higher than that of any pre-event air south of 65° N. Thus, if we postulate strict conservation of *PV*, no air from south of 65° could have crossed the gap and been mixed into the vortex. However, it is important to note that the size of the observed *PV* gap is close the effective resolution limit of the smoothed data used in our wind-field solution, so the signature of a hypothetical narrow filament of low-*PV* air penetrating the polar vortex would be blurred out into the relatively minor *PV* gap that we observe.

The second times series (Fig. 2.22) presents typical northern midwinter behavior. Here, the vortex shows minor distortions that vary from sol to sol and are mirrored by the exterior particle paths. The next time series (Fig. 2.23), from the"double-ring" period in the southern hemisphere, presents a much more con-

fusing picture. The particle paths have a no regular *PV* pattern to follow, and it's not clear that they have any connection to the *PV* field at all. Arcs of low *PV* air are present at all latitudes north of 75° south, and in no case is it clear how one sol's *PV* field deforms into the next.

Figure 2.24 shows the behavior of the northern polar vortex in its most compressed state during the planet-encircling dust storm of 2001. During these sols the center of the polar vortex is clearly offset by at least 5° latitude, and the direction of offset rotates slowing around the pole. As in the the other northern hemisphere time series, the particle paths and *PV*-field are offset together. Just as in (Fig. 2.21) there is a gap (or two) in the high-*PV* ring, but the *PV* in these gaps is high enough that no air from south of 70° N could have crossed it. Of course, the previously discussed caveat about the resolution of our data set applies to this situation as well.

2.3.4 *PV*-field breakdown events

In Figs. 2.25 – 2.31, we use time series of *PV* and trajectories as well as time series of temperature to explore the two major *PV*-field breakdown events.

In Figs. 2.25 – 2.28 we find the the time series bear out our earlier observation that the northern MY 26 breakdown actually involves a displacement, contraction, and strengthening of the high-*PV* region. The displacement might also be thought of as an asymmetric contraction, with the greatest contraction happening at 90° West. The longitude of greatest contraction remains fixed throughout the event, and the low-temperature region as well as the trajectories are displaced together with the *PV* field. For all three, the magnitude of the asymmetry is about 10° in latitude, meaning the the trajectories and contours are all 10° closer to the pole

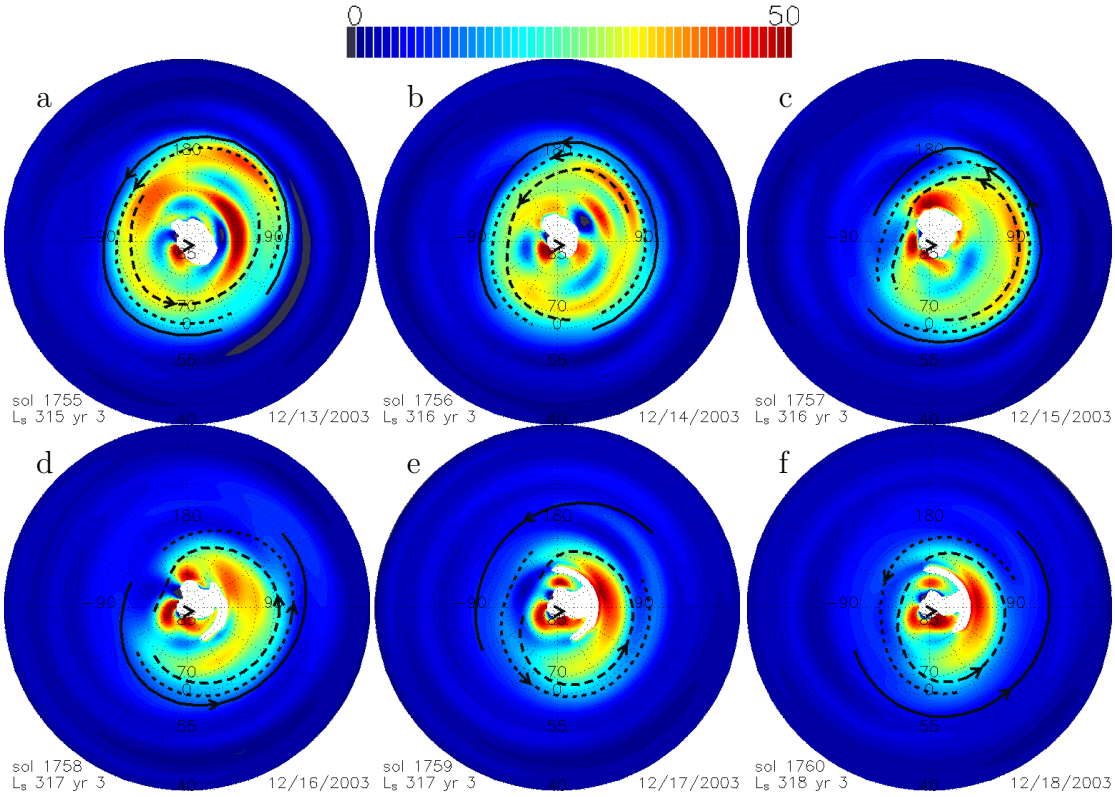

Figure 2.25: Time series of *PV* maps, sols 1755 – 1760 in the northern hemisphere, on the $\theta = 300$K surface, showing the onset of the *PV*-field breakdown discussed in the text. Otherwise the same as Fig. 2.21.

at 90° West than at 90° East.

The asymmetry of the temperature field is particularly important for understanding the apparently dramatic zonal-mean warming in Fig. 2.13. Despite the 15 K zonal mean temperature rise that peaks just before L_s 320°, the displaced cold core of the vortex remains at or below the saturation point all the way to L_s 321° sol 1765. Thus, we can infer that up until this point, although some latitudes and longitudes may be experiencing warming due to the dust storm distorting the polar vortex, the polar vortex itself has not been breached. However, on the following two sols, the trajectories, *PV*, and temperatures all point to the occurrence of a breach in the vortex. On sol 1766 a region of very low *PV* (< 0) near 0° East,

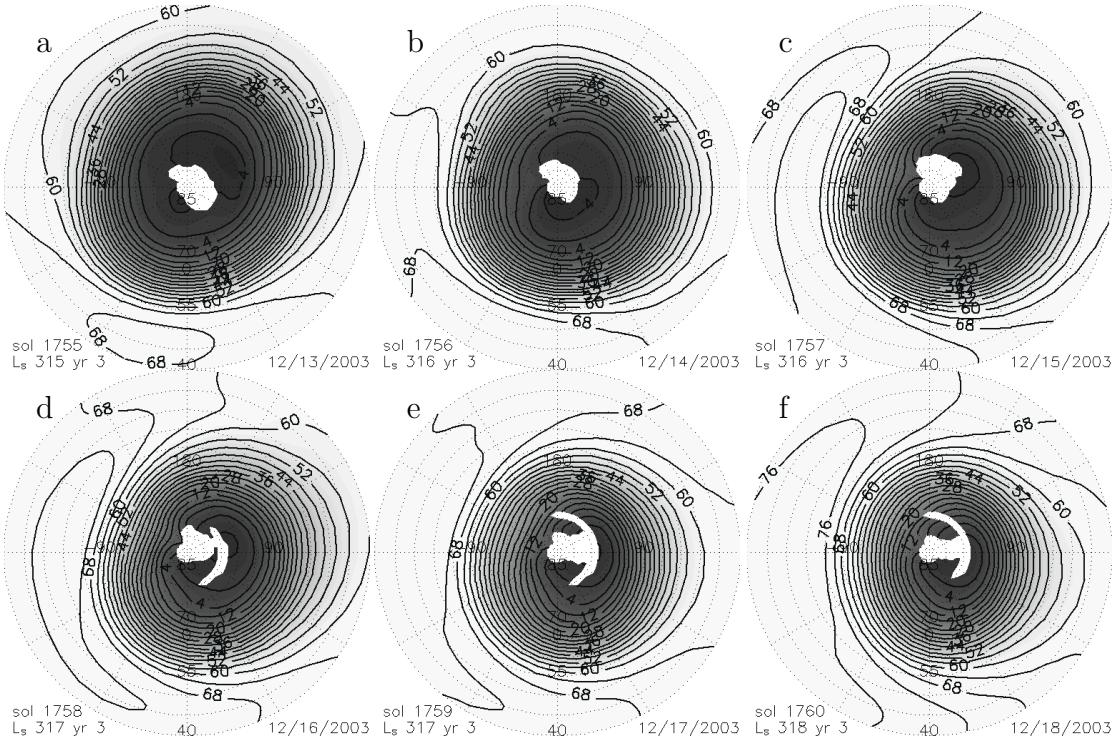

Figure 2.26: Maps of Temperature (K) expressed as $T - T_{\rm sat}$, for the same sols as in Fig. 2.25 (sols 1755 – 1760), at the 0.83 mb pressure level in the northern hemisphere.

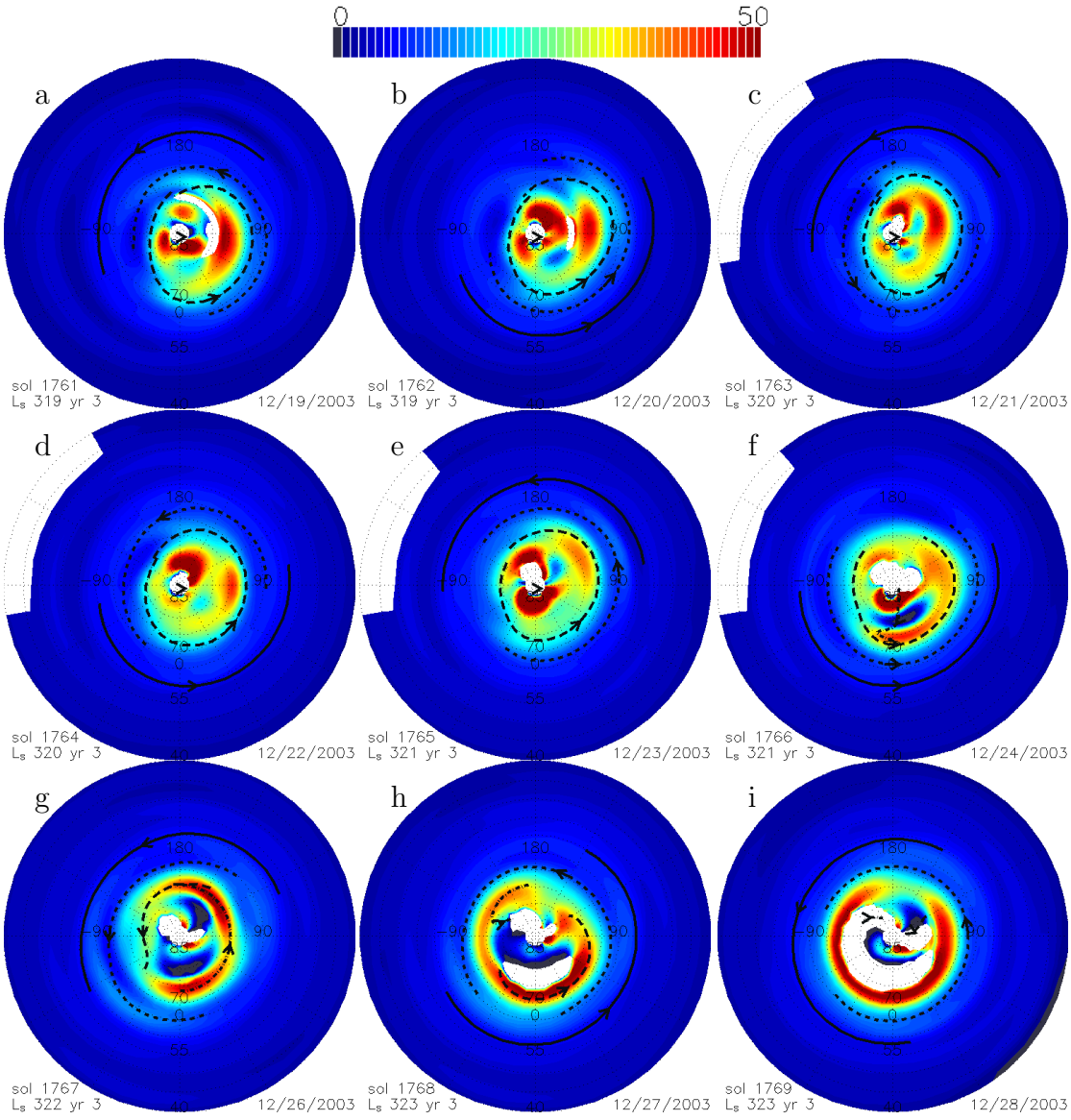

Figure 2.27: Time series of *PV* maps, sols 1761 – 1769 in the northern hemisphere, on the $\theta = 300$K surface, showing the decay of the *PV*-field breakdown event discussed in the text. Otherwise the same as Fig. 2.21.

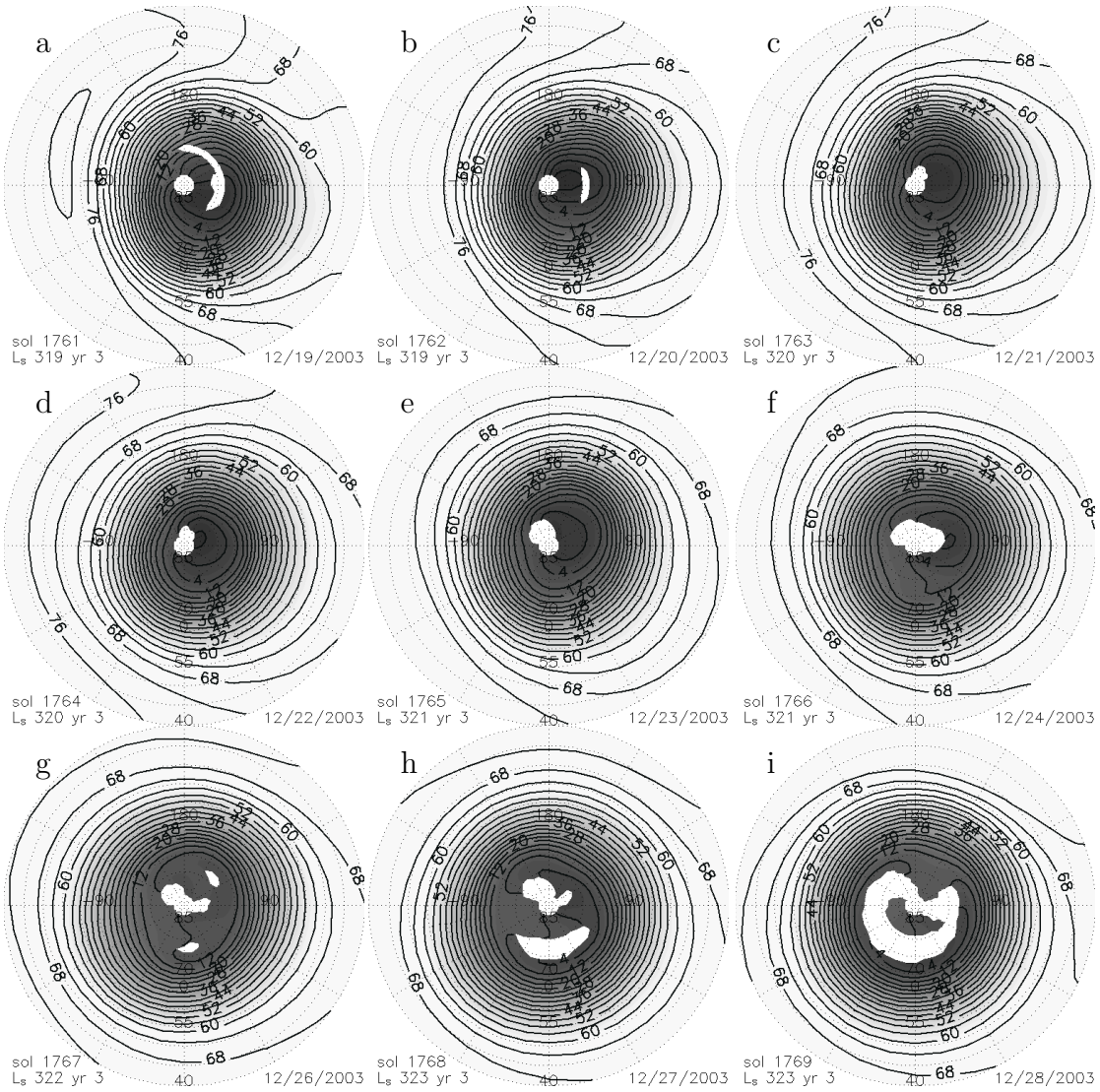

Figure 2.28: Maps of Temperature (K) expressed as $T - T_{\rm sat}$, for the same sols as in Fig. 2.27 (sols 1761 – 1769), at the 0.83 mb pressure level in the northern hemisphere.

75° North appears to have penetrated the high *PV* region, and easterly flow is observed in the same region. On this sol, cold temperatures ($T - T_{\rm sat} < 0$) are still observed inside the vortex at 90° East, but by the next, sol 1767, the entire vortex has warmed, with the minimum temperatures near $T - T_{\rm sat} = 10$. Immediately following this mixing event, the *PV* field returns to a more symmetric configuration, and polar temperatures start to drop.

In Figs. 2.29 – 2.31, we see that that southern MY 26 *PV*-field breakdown appears to be a prolonged period of warm temperatures and probably enhanced mixing, in contrast to the gradual distortion and sudden mixing of the northern event. However, the onset of the warm period is obscured by several sols of missing data, and so is not included in the time series that we present. Thus it is possible that the onset is just as abrupt despite the fact that the duration is much longer and the decay more gradual.

Since the southern *PV*-field is commonly highly disorganized, the *PV* field in Figs. 2.29 and 2.30 is not particularly remarkable, except that many of the features are readily traceable for several sols. The traceability of *PV*-field is likely a consequence of the fact that the wind speeds are very small for much of this time series. Other than the temperatures (polar warming with mid-latitude cooling as discussed earlier), the best evidence for enhanced mixing is the frequency with which trajectories at all latitudes reverse directions. In Fig. 2.29, one reverses directions four times, essentially circling a point near 0° East, 60° South rather than the pole. During the decay of the high-temperature period, the maximum in the *PV* field does not generally increase, instead the area covered by higher *PV* values gradually increases as the polar temperatures drop and the wind speeds increase. Although the temperature field during the warm period is fairly symmet-

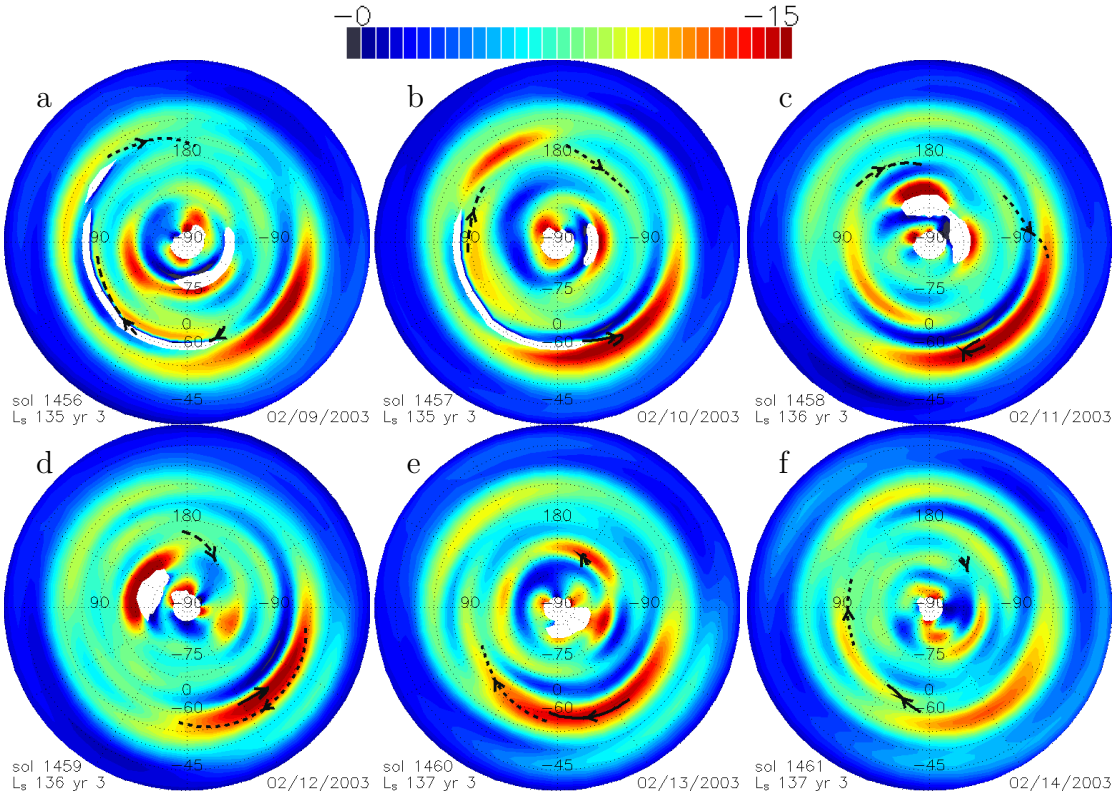

Figure 2.29: Time series of *PV* maps, sols 1456 – 1461 in the southern hemisphere, on the $\theta = 280$K surface, showing the peak of the south-pole polar vortex breakdown discussed in the text. Otherwise the same as Fig. 2.21.

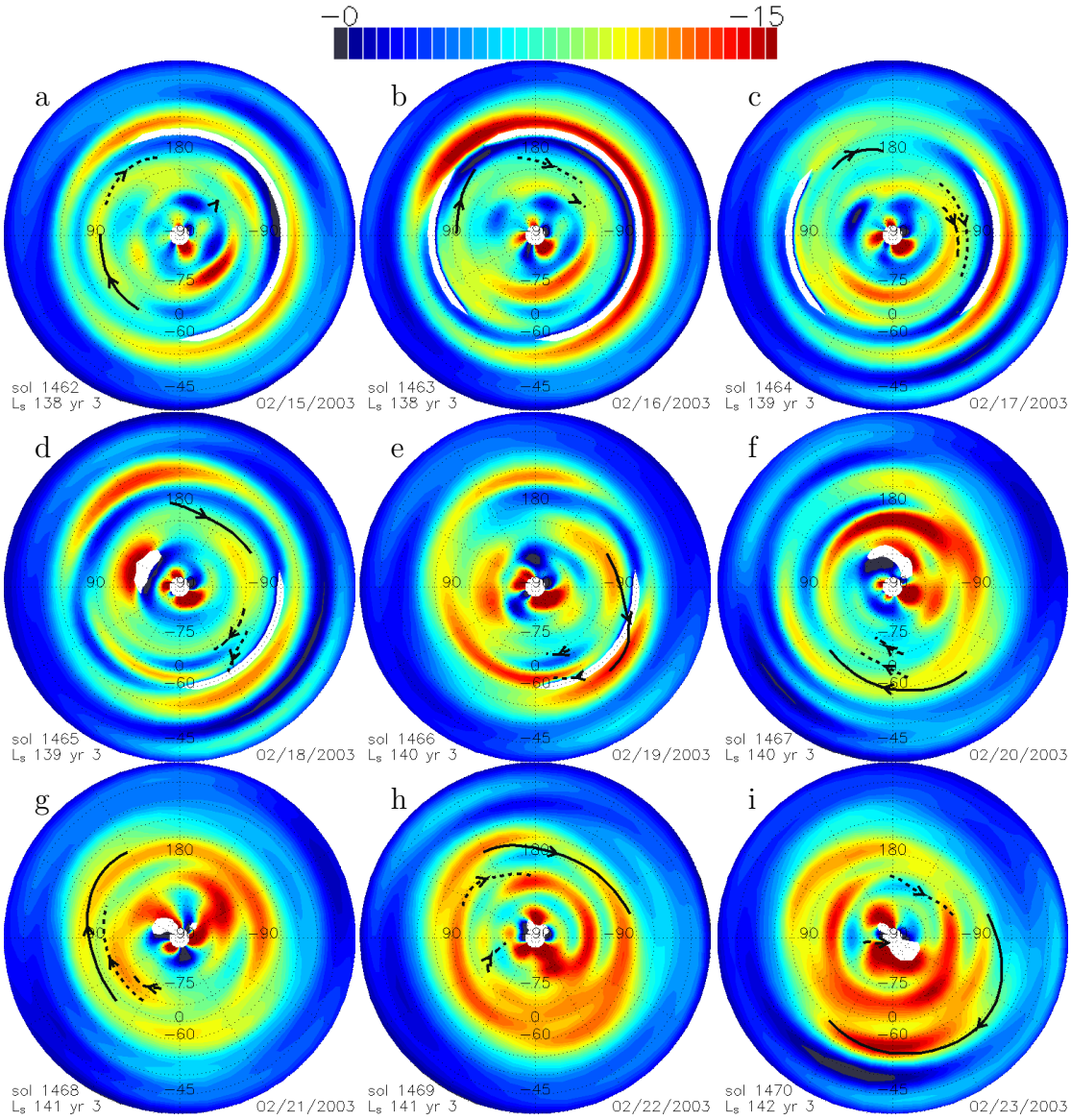

Figure 2.30: Time series of *PV* maps, sols 1462 – 1470 in the southern hemisphere, on the $\theta = 280$K surface, showing the decay of the south-pole polar vortex breakdown discussed in the text. Otherwise the same as Fig. 2.21.

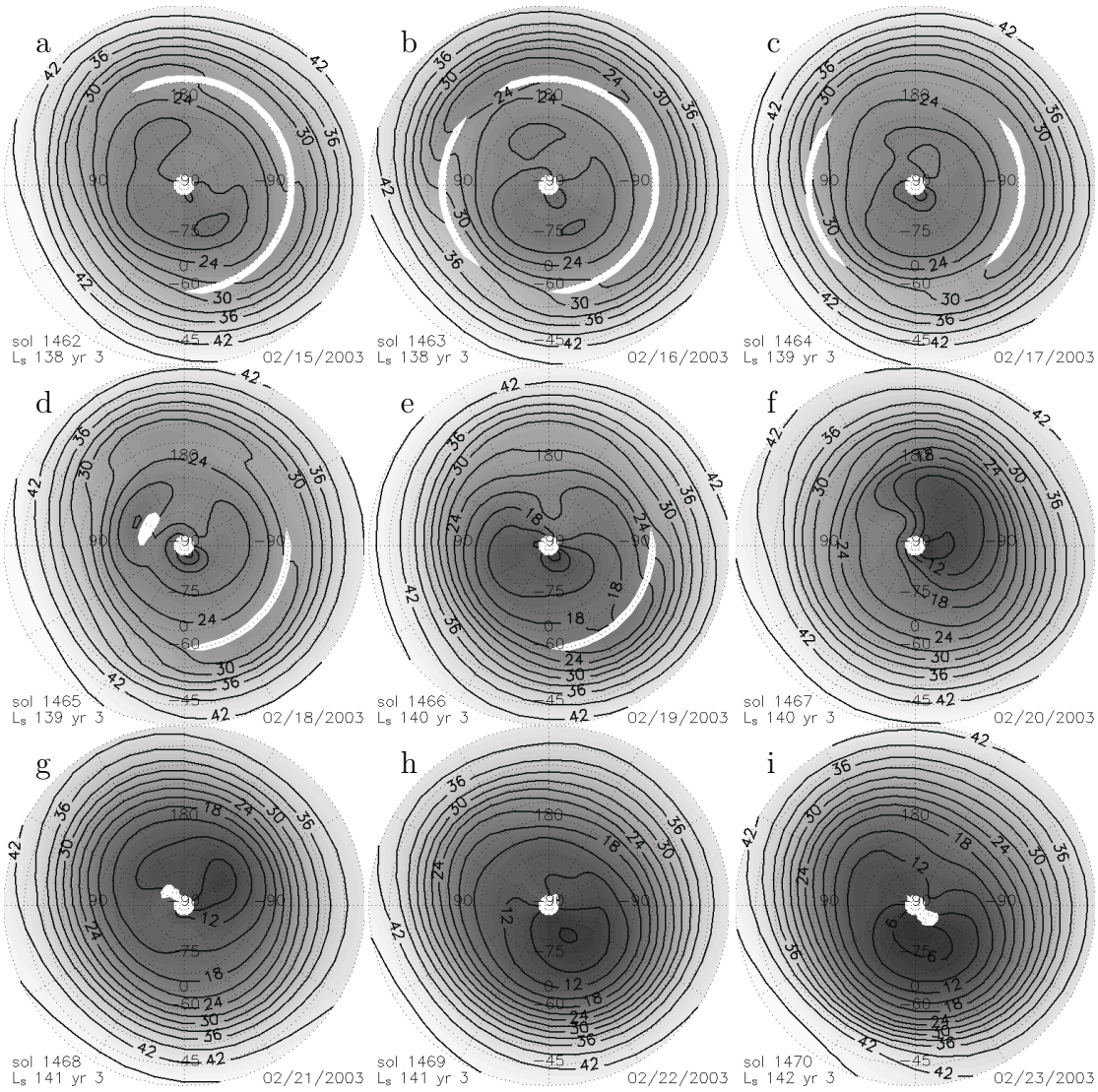

Figure 2.31: Maps of Temperature (K) expressed as $T - T_{\rm sat}$, for the same sols as in Fig. 2.30 (sols 1462 – 1470), at the 0.83 mb pressure level in the southern hemisphere.

ric, the temperature decrease associated with the decay is not initially symmetric, suggesting that some dynamical process is involved.

2.4 Discussion

2.4.1 *PV* generation

In order to understand the *PV* field, it is important to understand the *rate* at which *PV* is created and destroyed by diabatic processes. The timescale for *PV* generation will tell us over what length of time *PV* is useful as a tracer. More importantly, the *PV* generation rate is also the *maximum* rate at which mixing can occur across *PV* contours on a θ-surface. The time-rate-of-change of *PV* at a given grid point on a θ-surface is given by [Andrews et al., 1987]:

$$\frac{\partial PV}{\partial t} + u\,\frac{1}{a\cos\phi}\frac{\partial PV}{\partial \lambda} + v\,\frac{1}{a}\frac{\partial PV}{\partial \phi} = -\frac{1}{\sigma}\,\frac{1}{a\cos\phi}\frac{\partial Q}{\partial \lambda}\,\frac{\partial v}{\partial \theta} + \frac{1}{\sigma}\,\frac{1}{a}\frac{\partial Q}{\partial \phi}\,\frac{\partial u}{\partial \theta} + PV\,\frac{\partial Q}{\partial \theta} - Q\,\frac{\partial PV}{\partial \theta}\;; \tag{2.21}$$

where

$$Q = \frac{\partial \theta}{\partial t}\;, \tag{2.22}$$

and

$$\sigma = -g_0\,\frac{\partial p}{\partial \theta}\;. \tag{2.23}$$

A detailed calculation of the heating rate Q is beyond the scope of this work. Therefore, in order to estimate the order of magnitude of the *PV* generation rate, we turn to published general circulation model results by Haberle et al. [1993]. We take Q from their Fig. 25b (northern winter solstice, dust optical depth 0.3), and all other quantities from our gridded data for MY 24 northern winter solstice period (sol 344). We then interpolate all quantities onto θ-surfaces and apply Eq. 2.21,

Table 2.3: Diabatic Contributions to *PV* generation as fraction of total *PV*

$\frac{1}{\sigma}\frac{1}{a}\frac{\partial Q}{\partial \phi}\frac{\partial u}{\partial \theta}$ / *PV*	$PV\frac{\partial Q}{\partial \theta}$ / *PV*	$-Q\frac{\partial PV}{\partial \theta}$ / *PV*	Total
0.00015	-0.047	0.129	0.082

Mean value of terms from Eq. 2.21 over the region 68°–71° north, 110°–220° east. Units are sol^{-1}.

evaluating derivatives according the conventions described in section 2. Note that we do not attempt to include the effects of latent heat release and atmospheric mass loss associated with CO_2 condensation.

The second and third terms on the left-hand side of Eq. 2.21 merely represent the adiabatic advection of *PV* on the θ-surface. Thus, in order to understand the extent to which the convenient assumption of *PV* conservation following the flow is valid, we need only to evaluate the diabatic terms on the right-hand side of Eq. 2.21. Since the model results presented by Haberle et al. [1993] are zonal averages, we neglect terms involving $\partial/\partial\lambda$.

We find that the diabatic component of $\frac{\partial PV/\partial t}{PV}$ is greatest where *PV* is greatest. Table 2.3 shows the mean values of the terms that make up the diabatic component over a portion of the *PV*-maximum region. (Since our *PV* field is quite symmetric at northern winter solstice, and since Haberle's Q values are zonally averaged, our estimates of these diabatic terms are nearly independent of longitude.) In this region, *PV* is being diabatically generated at a rate of about 10% per sol; in other words, *PV* is generated with a time constant of about 10 sols. It is important to realize that this time scale may or may not be longer when the absolute values of *PV* are larger, since larger *PV* might well be associated with larger Q.

The dominant contribution to *PV* generation is the advection of *PV* *across* θ-surfaces by (in this case, radiative) cooling, Q. In other words, *PV* increases rapidly

with height, since it depends on ρ^{-1}, and so downward motion in the θ coordinate system carries high-*PV* air to lower levels. The vertical gradient in Q acts to somewhat counteract this effect, and horizontal gradients in Q are essentially irrelevant. $Q\,\frac{\partial PV}{\partial \theta}$ as the dominant *PV* generation mechanism is interesting in part because the same mechanism is responsible for the build-up of *PV* at the winter pole on Earth [Nash et al., 1996]. It's also interesting because the Hadley circulation is the likely source of the diabatically descending air represented by Q, which means that the intensity of the polar vortex *PV*-field is likely to be directly related to the strength of the Hadley circulation.

2.4.2 Mixing

Since the *PV*-map time series clearly show that dynamical timescales are on the order of one sol, it is reasonable to suppose that *PV* is to good approximation a tracer of atmospheric motion, and that rapid changes in the *PV* configuration, such as the spiral event of Fig. 2.21, are likely a result of dynamical processes rather than a simple response to changes in heating. Furthermore, no disturbance with a time scale of less than ten sols can mix air into the polar region without producing a noticeable rearrangement of the *PV* field. Thus, when the vortex is well organized, as in early–middle northern winter, it is clear that mixing is inhibited.

On time scales comparable to, or longer, than 10 sols, the same processes that generate *PV* may in principle mix atmospheric constituents across *PV* contours and thus into the high-*PV* region. Thus, for example, dust could be transported into the high-*PV* region via the radiative cooling branch of the Hadley circulation, but not by a primarily adiabatic process like a stationary wave. Our simple estimate of how Q affects *PV* does not, however, give any indication of how *PV* might be destroyed.

So once the aforementioned dust-bearing air reaches the high-*PV* region, it is not at all clear how, or how rapidly, or by what process, it could then *lose PV* so as to move into the low-*PV* core of the polar vortex. The ten sol timescale is therefore a *lower limit* on the timescale for mixing under organized-vortex conditions.

Applying a detailed radiative transfer calculation to our data set would help answer the *PV*-loss question. Even in the time period when the polar vortex is in its most organized phase, the inner edge of the high-*PV* region sometimes (*cf.* Figs. 2.21, 2.24, 2.27) penetrates all the way to the pole, and so dynamical mixing might also be important. Latent heat is, of course, another potential mixing driver, since the atmosphere is generally close to or past the condensation point just interior to the high-*PV* region.

When the vortex is poorly organized, as is common in southern winter and in late northern winter, our data set is consistent with, but does prove, that dynamical mixing is occurring. We are limited by the fact that our spatial and temporal resolution are probably too poor to resolve the processes involved. Since the *PV* field is annular, the mere existence of low-*PV* air at a high latitude does not in general provide evidence of mixing as it would in the terrestrial stratosphere. To conclusively demonstrate mixing from the *PV* field alone, we would therefore need to be able to follow the migration of a low-*PV* feature from the mid-latitudes towards the pole, but the usually rapid evolution of *PV* features makes this impossible. An event in which the high-*PV* boundary between mid-latitude and polar air was breached would also imply mixing, but since our wind solution has a resolution of no better than 50° in longitude and 5° in latitude, only a very large breach would be clearly observable. Smaller breaches would appear smoothed out into the suggestive but not conclusive low spots that we commonly observe in the "disorganized" periods.

However, since there is a clear contrast between the organized and disorganized time periods, we can still draw the general conclusion that dynamical mixing processes such as wave-breaking are much more likely to be important, *i.e.*, operate at a higher rate, in southern polar winter than in early – middle northern polar winter. If the dynamical mixing time scale in the southern hemisphere turns out to be shorter or comparable to the time scale for diabatically driven mixing, then the southern winter will have a much higher mixing rate overall.

In addition to these periods of enhanced but not clearly resolved dynamical mixing, we have identified two large-scale mixing events that, as previously discussed, are discernible in temperature and trajectories as well as in *PV*. For the southern winter event, the very rapid temperature rise implies that some dynamical process was the cause, but due to a gap in the TES data set, we cannot say anything more about what triggered it. For the northern winter warming event, however, the apparent penetration of the vortex by a low *PV* airmass just prior to the warming of the vortex core strongly suggests a planetary wave breaking event similar to a terrestrial sudden stratospheric warming.

2.4.3 Connection between the polar vortex and dust opacity

Some of the correlation between northern polar vortex behavior and low-latitude dust opacity may be explainable in light of, 1) the previously discussed dominant role of radiative cooling in generating potential vorticity, and 2) simulations by Wilson [1997] that indicate an expanded Hadley circulation in response to dust. If we suppose that the zone of radiative cooling that generates *PV* is in fact the descending branch of the Hadley circulation, then, as the expansion of the Hadley

circulation pushes the descending branch northward, the high-*PV* zone must move northward along with it, just as we observe. In order to verify this hypothesis, it will be necessary to, 1) evaluate Q directly from our data set in order to understand the precise nature of *PV* generation, and 2) calculate *PV* from modeling results that show expansion of the Hadley circulation.

However, the behavior of the polar vortex during the three high-dust periods that we have studied indicates that the response of the polar vortex to dust events is more complicated than a simple Hadley cell expansion. First of all, all three (Figs. 2.21, 2.24, 2.25) are associated with large ($\sim 10^\circ$ latitude) asymmetric perturbations to the *PV* maximum. The asymmetry may simply be a reflection of asymmetric dust loading. Nevertheless, since these perturbations occur on two-sol timescales, they are too rapid to be explained by diabatic *PV* generation, and so must be associated with some dynamical phenomenon, even if dust opacity was the root cause. Furthermore, the response of the vortex to to a dust storm is not simply related to the intensity of the storm. The MY 25 (August 2001) dust storm produced substantially higher dust opacity than the MY 26 (December 2003) storm, but the MY 26 storm produced a much more significant disturbance in the polar vortex *PV* field. Finally, although the MY 26 storm's warming event is initially consistent with a Wilson [1997]-style Hadley cell expansion, the core of the polar vortex ultimately warms only in response to a planetary-wave-breaking event.

2.4.4 Polar vortex behaviors unrelated to dust

None of the differences between the two southern winter seasons appear to have any connection with dust opacity. The TES data set does not point to any explanation for these differences. Modeling work to reproduce these interannual

differences could lead to understanding of an important interannual variability in mixing rates. Bridger et al. [2003] have conducted multi-year GCM simulations that show significant interannual variability in poleward eddy heat transport, but they do not discuss *PV*, and report only a small amount of interannual variability in the southern hemisphere.

Although we can't explain why the south pole warming event of MY 26 occured, we can say that this event is fundamentally different from the northern MY 26 event because it involves no Hadley cell expansion, as evidenced by the fact that the *PV* field simply diminishes instead of migrating poleward, and because the the vortex core remains warm for much longer. However, the reversal in wind direction shown by one trajectory in Fig. 2.27 is reminiscent of a terrestrial stratospheric warming, as is the large amount of low-*PV* air near the pole. Thus, there is a good chance that this event was also caused by a terrestrial-like planetary-wave-breaking episode.

2.4.5 Instabilities

Both the symmetric ($f\,PV < 0$) and barotropic (*PV* has a local maximum or minimum) instability criteria are frequently satisfied in this data set. Most of the symmetric instability is found in a broad mid-latitude swath above the 1 mb level in northern winter. The only examples of significant symmetric instability in the polar region occur during the two major mixing events.

Many of the barotropically unstable *PV* features are transient. However, the basic state of the winter polar atmosphere on Mars appears to be annular in *PV*, and so inherently satisfies the barotropic instability criterion over the entire vertical range of our data set. However, a local *PV* maximum or minimum is merely a *necessary* condition for instability. More importantly, this instability criterion

is derived by linearization about a quasi-geostrophic basic state [Andrews et al., 1987]. Since the Mars polar vortex generally has a Rossby number (R) on the order of unity, it is far from the quasi-geostrophic regime, and so the standard barotropic instability theory may not apply. Clearly an analytical treatment of Rossby waves for the $R \sim 1$ case will be necessary to understand Mars polar vortex dynamics. Note that barotropically unstable zonal mean flow *is* observed on Earth, in small regions just inside and just outside of the southern winter polar vortex [Hartmann, 1983; Andrews et al., 1987], but the wind speeds there fall well within the small R regime. Barnes et al. [1993] note in passing that the meridional potential vorticity gradient changes sign on the flanks of the polar vortex jet in their Mars GCM simulations. However, their model exhibits these sign changes only at high altitudes and only in the northern hemisphere, and they do not discuss the stability implications of this phenomenon.

2.4.6 CO_2 saturation

The coverage of MGS radio occultation profiles (see, for example, `http://nova.stanford.edu/projects/mgs/coverage.html` and Hinson et al. [1999]) is very sparse compared with the TES temperature measurements that we rely on, and winter pole measurements turn out to be extremely rare. However, it may be possible to use the few winter polar observations that do exist to check our supersaturation results. The MGS radio occultation temperature-pressure profile data set is described by Hinson et al. [1999].

The equatorward boundary of the region where we measure saturated temperatures does coincide, at both poles, with the equatorward boundary of the region where the Mars Orbiter Laser Altimeter (MOLA) detects high-opacity clouds [Co-

laprete et al., 2003]. Furthermore, Colaprete et al.'s maps of MOLA cloud returns show a higher incidence of clouds within a ring roughly coincident with the zone of highest saturation that we measure just inside the *PV* maximum. Thus, regardless of the accuracy of the temperature measurements, the regions of saturation that we identify interior to the polar vortex are probably correct in the sense that they reflect zones of active CO_2 condensation. If the temperature measurements are in error, then high-opacity CO_2 cloud should be considered as a possible cause of that error, which would mean that spuriously low temperatures are indirectly valuable as an indicator of opacity.

Assuming that the saturated regions are correctly identified, we can draw two important conclusions. First, since the high-*PV* zone and the saturation zone move together, with the saturation zone staying just inside the high-*PV* zone, the high-*PV* zone may be controlling the location of the saturation zone by acting as the poleward boundary of heat transport. The converse — that latent heat release in the saturated zone would block the diabatic cooling that generates *PV*, and so define the poleward boundary of the high-*PV* region — seems less likely given that the annular *PV* field persists even at altitudes above where the saturation takes place. Second, the MY 26 southern winter polar vortex probably had fewer clouds and less atmospheric CO_2 condensation than the preceding year. Less CO_2 "snow" has implications for the surface properties of the polar cap when it is revealed in spring.

2.5 Conclusions

1. The Mars polar vortex has a Rossby number of order unity and an annular *PV* field. The stability of this configuration needs to be investigated.

2. The *PV* field of the polar vortex approximately defines the boundary of the CO_2 saturation zone.

3. The northern hemisphere polar vortex is more organized and therefore probably a stronger barrier to mixing and heat transport than the southern hemisphere polar vortex. Radiative heating/cooling estimates from the TES data set, as well as numerical models, are needed to verify and quantify this result.

4. Radiative cooling associated with the Hadley circulation is probably the main generator of *PV*. Radiative heating/cooling estimates from the TES data set, as well as numerical models, are needed to verify and quantify this result.

5. Increased low-latitude dust loading leads to contraction of the polar vortex, but in general has no impact on temperatures in the core of the vortex. The onset of the contraction in response to dust storms is too fast to be explained by a purely diabatic process.

6. The southern polar vortex shows substantial interannual variability that is not related to dust loading.

7. The MY 26 southern polar vortex experienced a 10 K warming event that is associated with a breakdown of the *PV* field.

8. The MY 26 northern polar vortex, during the December 2003 dust storm, experienced a planetary-wave-breaking event that was followed immediately by a 10 K warming in the core of the vortex.

CHAPTER 3

THEMIS-VIS OBSERVATIONS OF CLOUDS IN THE MARTIAN MESOSPHERE

3.1 Introduction

The Mars Odyssey (ODY) spacecraft's Thermal Emission Imaging System (THEMIS) [Christensen et al., 2004] has been conducting mapping operations from Mars orbit since February 2002. Although the THEMIS investigation is directed primarily towards surface geology, [*e.g.*, Christensen et al., 2005; Bandfield et al., 2004; Christensen et al., 2003; Titus et al., 2003] it routinely monitors aerosol opacities [Smith et al., 2003] with its multispectral infrared detector (THEMIS-IR) and images clouds [Richardson et al., 2006] with its visible band subsystem (THEMIS-VIS).

For a subset of the THEMIS-VIS cloud images, it is possible to discern movement of the cloud features relative to the surface. This apparent motion yields information about the altitude of the cloud, via parallax, as well information about the actual velocity of the cloud. This chapter focuses on those cloud features with the largest apparent motion, which are those at mesopheric [as defined by Zurek, 1992] altitudes. At these altitudes, the apparent motion is dominated by parallax, but zonal velocities are also measurable.

Mesospheric clouds, and their advection by mesospheric winds, are key indicators of mesospheric dynamics, are an important test for models of cloud microphysics and the martian general circulation, and may provide insights into the possibly analogous terrestrial phenomenon of noctilucent clouds. They might also be important for the radiative budget and chemistry of the mesosphere. The prevail-

ing explanation for the observed reversal in the mesosphere of the pole-to-equator temperature gradient [*e.g.*, Smith et al., 2001] on both Mars and Earth is that upward propagating gravity waves break in the mesosphere, thereby transporting momentum from the surface and exerting a drag on mesospheric winds. [Holton, 1983; Jaquin, 1989; Joshi et al., 1995] Condensate cloud morphology may provide direct evidence of gravity waves, and the wind field traced by cloud motion is of course sensitive to the amount of drag. Wave drag is a crucial boundary condition for general circulation models. [*e.g.*, Joshi et al., 1995]

In order for clouds to form, aerosol particles must be supported by updrafts or eddy diffusion, or condensation must be rapid. Models of mesospheric cloud formation [Glandorf et al., 2002; Colaprete and Toon, 2003] require heterogeneous nucleation, unless the mesosphere is much colder than expected, and so dust grains of sufficient size must be transported upward from the surface, imposing requirements on eddy diffusion and/or vertical velocities for the entire atmosphere. Similar problems exist in the study of terrestrial noctilucent clouds (most recently reviewed by Kokhanovsky [2005]), which are observed in twilight and found in the summer polar mesosphere. Hunten et al. [1980] suggested meteors and micrometeors as a possible solution to the terrestrial nuclei source problem. The most significant mystery surrounding terrestrial mesospheric clouds, however, is that reports of noctilucent clouds seem to begin in the industrial era, and their frequency and latitudinal extent have shown a secular upward trend [Klostermeyer, 2002].

Jaquin [1989] suggests that even the small vertical optical depth of the mesospheric haze might be significant for the radiation budget of the poles, once the slant-path optical depth is taken into account. Clearly, however, mesospheric clouds have substantial impact on the optical depth of the mesosphere itself, and

thus may significantly influence the mesosphere's ability to absorb and emit radiation. Mesospheric clouds may also impact the chemistry of the mesosphere. Atreya and Blamont [1990] invoke heterogeneous chemical reactions, permitted by the presence of the aerosols, to explain an observed enhancement in the rate of CO recycling back to CO_2. They suggest that a persistent depletion of CO relative to CO_2 might decrease the long term loss of CO_2 to space.

Jaquin et al. [Jaquin et al., 1986; Jaquin, 1988, 1989] and Clancy et al. [2006] present surveys of mesospheric aerosols observed on the limb by the Viking orbiters and by Mars Global Surveyor, respectively. These aerosols are observed in the form of a "detatched haze," for which profiles of brightness above the martian limb show at least one distinct peak at an altitude of 50 – 90 km. The detached hazes have been detected only in visible wavelengths — by the Viking Orbiter, the Mars Orbiter Wide-Angle Camera, and the TES solar-band bolometer — and have vertical optical depths on the order of 0.01. Some images reveal multiple peaks in the brightness profiles, which could be either multiple vertical layers or horizontally discrete clouds in the foreground or background. The mesospheric hazes are not detected by the TES infrared channels, which Clancy et al. [2006] suggest implies an effective particle radius of less than 0.5 μm.

Neither survey detects any detached haze above 50 km during the aphelion period, L_s 60° – 100°. Jaquin [1989] reports that the maximum elevation of limb aerosols oscillates in phase with the annual variation of solar insolation, with the minimum below 50km during the aphelion period. Clancy et al. [2006] report that the detached hazes are most prevalent immediately before and immediately after the aphelion period, and are concentrated within two longitude bands: 240° – 310° east and 340° – 20° east. These two surveys do not provide a clear answer about

the prevalence of high-latitude mesospheric detached hazes. While Clancy et al. [2006] describe the detached hazes as largely confined to within 15° of the equator, Jaquin [1988, 1989] report a few examples in the 50° – 60° south range near L_s 300°, as well as in the 40° – 50° north range near L_s 330°.

Spectroscopic solar occultation measurements performed by the Phobos 2 spacecraft [Chassefière et al., 1992] provide a much smaller data set and more limited temporal coverage, but offer more detailed information about the limb aerosols. All of the Phobos 2 detached hazes were between 45 and 60 km altitude, have a vertical thickness of 3 - 6 km, and a vertical optical depth of less then 0.03. Chassefière et al. [1992] report a diversity of particle sizes, ranging from an effective particle radius of 0.15±0.1 μm in one measurement to 0.9±0.2 μm in another.

Horizontally resolved cloud features (as opposed to limb hazes) at mesospheric altitude have been previously reported on only one occasion, and the altitude of those clouds is in dispute. On sol 39 of Mars Pathfinder lander operations, "discrete blue clouds" [Smith et al., 1997] were observed by the Imager for Mars Pathfinder (IMP) during a period from 100 to 35 minutes prior to sunrise. The visibility of these clouds before sunrise could in principle yield information about their altitude, but the IMP science team [Smith et al., 1997] explicitly ruled out the possibility that these were directly illuminated mesospheric clouds, arguing instead that they were indirectly illuminated clouds at lower altitudes, and thus most likely composed of water ice. Clancy and Sandor [1998] describe the same features as "discrete linear clouds" with "wave-like structure," and argue that they were directly illuminated mesospheric clouds at 80 – 100 km altitude and composed of carbon dioxide ice. Their preferred model for the aerosol is 0.1–0.3 μm effective radius dust particles with number densities of 20–1000 cm^{-3} and a vertical optical

depth of 0.1, an order of magnitude greater than the detached limb hazes.

The composition of mesospheric aerosols has been controversial, in part because the behavior of martian mesospheric temperatures is not well understood. Earlier authors [Chassefière et al., 1992; Jaquin, 1988] ruled out the possibility of CO_2 ice on the basis of temperature measurements, and conclude that the aerosols must be composed of water ice. Clancy and Sandor [1998], however, review a larger set of historical data, including the Pathfinder Lander descent profile and ground-based submillimeter measurements, and argue that the martian mesosphere is cold enough in some seasons to allow for CO_2 condensation. In their view, the mean mesospheric profiles are at times within 10 K of the CO_2 frost point, which allows gravity waves to transiently drive the temperature below saturation as was seen in the Pathfinder descent ([Magalhães et al., 1999]). Modeling by Colaprete and Toon [2003] presents a complication. Colaprete and Toon [2003] calculate that CO_2 nucleation at the level of supersaturation seen in the Pathfinder profile requires dust particles of at least 0.5 μm in radius, which is obviously too large to form the 0.1–0.3 μm ice particles suggested by Clancy and Sandor [1998].

Previous estimates of mesospheric zonal (east-west) winds have been made either by applying a gradient-winds methodology to temperature soundings [*e.g.*, Smith et al., 2001], or by earth-based doppler techniques [Lellouch et al., 1991; Sonnabend et al., 2005]. The gradient method is not applicable near the equator, and it requires an assumption about the surface winds as a boundary condition. The resolution of the prerequisite temperature field is also typically fairly low — 30 degrees sampling interval in longitude and 1 scale height in altitude in the case of MGS-TES [Smith et al., 2001] — and so the wind estimate is inherently an average over large vertical and horizontal scales. The TES gradient wind estimates

are available only below 0.01 millibars (about 60 km altitude), but they have a key advantage in that they are available planet-wide in a continuous time series over a period of several martian years. The two earth-based doppler measurements provide a more direct estimate of mesospheric wind, but with FWHM beam sizes of 50% and 15% of the planet's diameter (Lellouch et al. [1991] and Sonnabend et al. [2005] respectively), they represent averages over a broad swath of the mesosphere. (The altitude resolution is much better, because both groups select spectral lines confined to 60 – 80 km altitude.) Lellouch et al. [1991] use a millimeter-wave carbon monoxide absorption band and assume a functional form for the shape of the expected easterly equatorial jet in order to compensate for their poor horizontal resolution. For their single measurement near northern winter solstice, they derive an easterly speed (*i.e.*, westward flow) of 160 ± 80 m/s for the core of a jet assumed to be centered at 20° south. Sonnabend et al. [2005] find a westward (easterly) zonal flow of 74 ± 22 m/s near 20° north in mid- northern winter using a mid-infrared carbon dioxide emission line.

In this chapter we describe ~ 50 measurements of altitude and velocity for mesospheric clouds that are horizontally resolved with a sampling interval of 70 meters per pixel or better. In order to place the measurements in context, we compare the images with daily global maps from Mars Global Surveyor's Mars Orbiter Wide Angle Camera (MOC-WA), and we compare measured winds results from a baseline general circulation model. Finally, we apply a radiative transfer model to a subset of these images in order to investigate the extent to which THEMIS-VIS can place constraints on the cloud aerosol properties.

3.2 Altitude-velocity measurement: theory

Apparent motion is detected as a side-effect of the THEMIS-VIS multi-spectral imaging scheme — the same location is imaged through different filters at different times, so that map-projecting the frames acquired by different filters onto the local surface will effectively co-align stationary surface features, but create a misalignment for any feature with apparent motion caused either by parallax or velocity. A very good approximation of this parallax motion is obtained by neglecting the curvature of the planet and taking the spacecraft trajectory to be parallel to the surface. The geometry of this situation is shown in Fig. 3.1. Suppose that THEMIS-VIS images a cloud feature through a certain filter while the ODY spacecraft is at position (1), and then images the cloud feature one second later through another filter while the spacecraft is at position (2). The apparent position of the cloud feature relative to the surface is the intersection with the surface of a line drawn from the spacecraft through the cloud feature, and the parallax motion is the distance between the intersections of the two lines from the two different spacecraft positions. Using similar triangles we find that the parallax motion,$v_{\rm app}$, is:

$$v_{\rm app} = -v_{\rm ODY}\frac{z_{\rm cloud}}{z_{\rm ODY} - z_{\rm cloud}} \quad , \tag{3.1}$$

and

$$\frac{dv_{\rm app}}{dz_{\rm cloud}} = -v_{\rm ODY}\frac{z_{\rm ODY}}{\left(z_{\rm ODY} - z_{\rm cloud}\right)^2} \quad . \tag{3.2}$$

With the ODY orbit altitude, $z_{\rm ODY}$, and ground-track velocity, $v_{\rm ODY}$, of approximately 400 km and 3000 m/s, a cloud at, for example, $z_{\rm cloud} = 50$km altitude has an apparent motion due to parallax of 430 m/s, and the derivative of the apparent motion is 10 m/s per km of altitude.

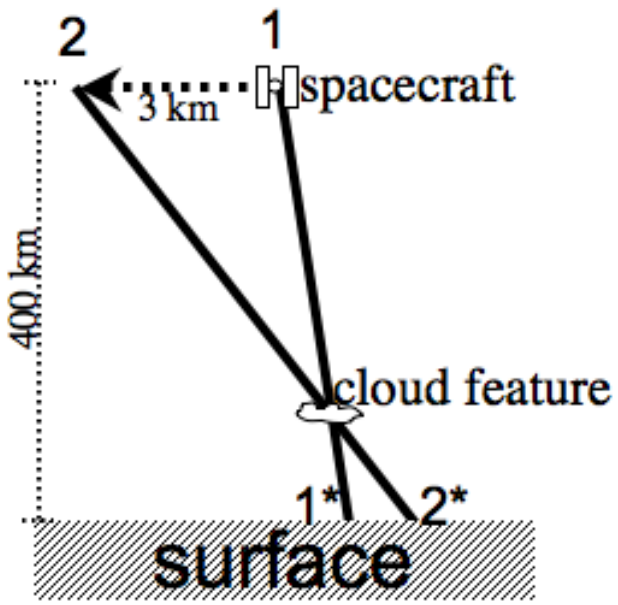

Figure 3.1: Schematic (not to scale) of parallax motion. When the spacecraft moves from position 1 to position 2 between exposures, the apparent position of the cloud feature relative to the surface moves from 1* to 2*.

The THEMIS-VIS pixel sampling interval is 0.045 milliradians [Christensen et al., 2004], which amounts to 18, 36, and 72 meters per pixel at the martian surface for the available 1x1, 2x2, and 4x4 binning modes, respectively. The narrow-band ($\sim$ 50 nm FWHM) THEMIS-VIS filters are bonded directly to the surface of the detector. Each filter has a field of view of 1024 (1x1 binned) pixels perpendicular to the ground-track of the spacecraft ("cross-track") by 192 (1x1 binned) pixels parallel to the ground-track of the spacecraft ("down-track"). The orbital motion of the spacecraft causes their fields of view to pass over surface features in the following order: 869 nm, 425 nm, 654 nm, 749 nm, 540 nm (the filters are labeled by their center wavelength). It takes slightly longer than 1 second for a surface feature to cross the down-track width of a filter, and so the delay between individual images in a THEMIS-VIS imaging sequence is set to a value close to 1

second. The interframe delay is exactly 1 second for all of the imaging sequences used in this chapter. See chapter 4 for a more detailed discussion of THEMIS-VIS operations.

We measure apparent motion using the 425 nm ("blue") and 540 nm ("green") filters, because these filters offer the largest change in apparent pixel position for any given apparent motion. (The 869 nm band is not usable due to severe stray light, see chapter 4.) The blue and green filters also offer the highest contrast for cloud features. Given the 1 second interframe delay, and the three filter widths between the blue and the green filters, most features observed in the blue filter are observed 3 seconds later in the green filter, and thus the 10 m/s per km of altitude of parallax motion translates to a 1.7, 0.83, or 0.42 pixels per km apparent motion in 1x1, 2x2, or 4x4 binned THEMIS-VIS images. Actual cloud velocity produces, of course, 0.17, 0.083, or 0.042 pixels per m/s of apparent motion. Thus, we expect that the precision of our measurements will not be much narrower than 20 m/s in velocity and 2 km in altitude for the 4x4 binned THEMIS-VIS sequences in which most of the high altitude clouds are detected. We also expect that, at mesospheric altitudes, the velocity motion will be much smaller than the parallax motion.

Since the parallax apparent motion is parallel to the ground track of the spacecraft (and in the opposite direction), the component of velocity parallel to the spacecraft ground track can not be distinguished from parallax motion. Since the parallax motion is the dominant effect, this means that the ground-track-parallel velocity component is essentially undetectable, and that the presence of a ground-track parallel velocity component introduces a small error in altitude which is given by equation 3.2. The Odyssey spacecraft is in a polar orbit, moving southward on the afternoon side of the orbit where all observations discussed in this chapter

were acquired. Thus, a 50 km high cloud feature moving southward at 50 m/s would have a measured altitude of 45 km. Meridional (north-south) winds in the martian mesosphere have never been measured, but the Joshi et al. [1995] model of gravity wave drag in the martian mesosphere leads to time-averaged meridional winds of less than 50 m/s for the regions probed by our measurements. This model calculation was for perihelion, the time of greatest meridional flow, and so a 50 m/s meridional wind can be considered an extreme case. Due to the rotation of Mars and the small inclination of the Odyssey orbit plane, the ground track is not actually due south — its heading is near 190° at mid- to low altitudes — but this introduces only a trivial (less than 2%) correction to the relationship between meridional velocity and altitude implied by equation 3.2.

3.2.1 Altitude-velocity measurement: method

In principle, we could solve for any one velocity component that is not parallel or nearly parallel to the ground track. We choose the zonal velocity simple because meridional and zonal velocities are the standard basis for describing horizontal atmospheric winds, and the meridional velocity component is unresolvable as previously discussed. Fortunately, the zonal velocity is normally the dominant wind component in the atmosphere (except near the planet surface).

Figures 3.2 – 3.5 illustrate our method using a particular measurement from THEMIS-VIS image sequence V04573003 as an example. Figure 3.2 shows a contour plot of the correlation of the blue filter image with the green filter image as a function of zonal (eastward) velocity and altitude above the Mars reference ellipsoid. Figure 3.3 shows the region of the V04573003 Reduced Data Record (RDR) used to calculate the correlation (described below), which we will henceforth refer

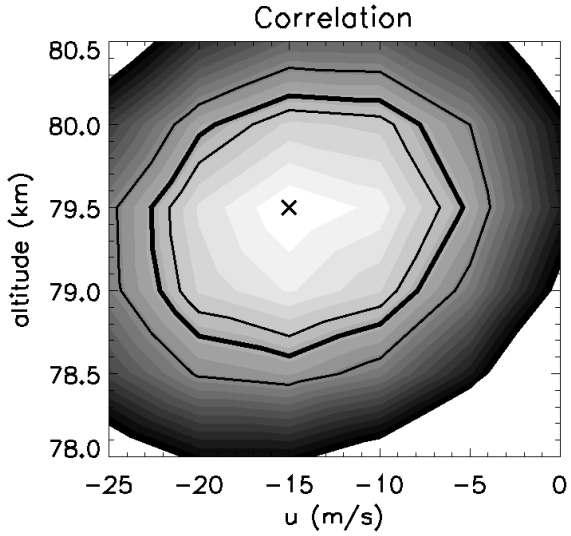

Figure 3.2: Correlation as a function of altitude and zonal velocity. The shaded contours show the difference between the correlation and the peak correlation (indicated by "X"), with contours at intervals of 0.25 in units of the standard error of the correlation. The black lines show a test of the hypothesis that the correlation at a given grid point is not less than the maximum correlation. The thick black line shows the 5% level for this hypothesis, which defines our confidence intervals. The outer and inner thin black lines show the 1% and 10% levels, respectively.

to as the correlation region of interest ("correlation-ROI"). Figure 3.4 shows the blue filter projected onto the green filter correlation-ROI at the best fit (maximum correlation) altitude and velocity, and Figure 3.5 shows the same projection using the altitude and velocity (zero by definition) of the local surface.

All of our measurements are performed using THEMIS-VIS RDRs, which are available from NASA's Planetary Data System. Each RDR contains all of the filter images from all of the exposures in a THEMIS-VIS imaging sequence. Each plane of the RDR stores all the individual filter images for a given filter, arranged in temporal order. We refer to these individual filter images as "framelets". Each framelet is a 1024 by 192 (or 512 by 96 or 256 by 48 for 2x2 or 4x4 binning) array of radiance values for a single filter in a single exposure. (See chapter 4 for the

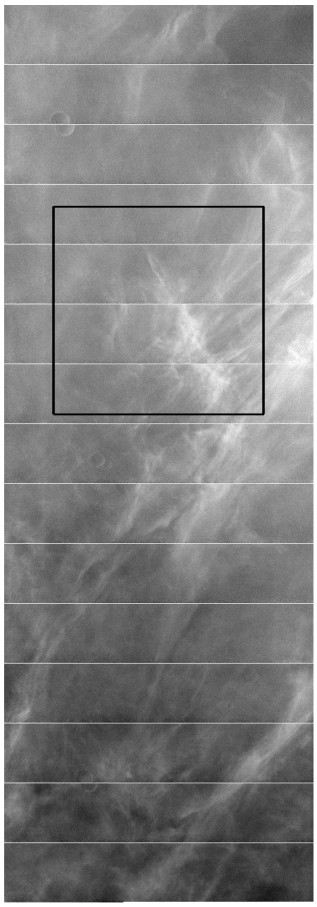

Figure 3.3: The V04573003c ROI, which was used to generate the correlation map shown in Figure 3.2.

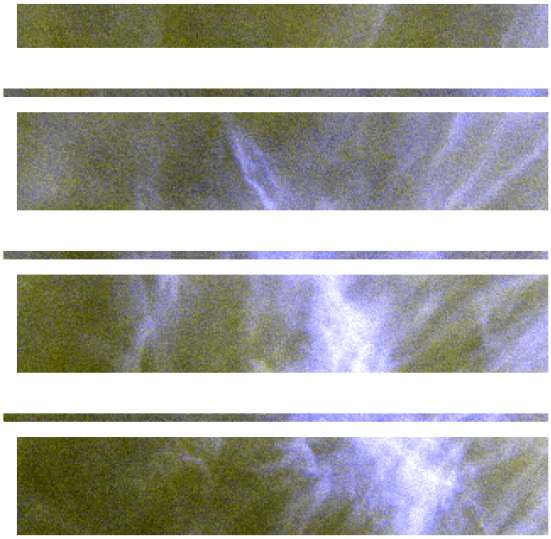

Figure 3.4: Co-alignment of the green and blue filters at the best-fit altitude and velocity solution for the V04573003c ROI. The blue filter has been projected onto the green filter RDR as described in the text. In this figure, regions of the ROI that are invalid for calculating the correlation according to the criteria described in the text are left blank. The red and green channels of this RGB figure are controlled by the green filter ROI image, and the blue channel by the projected blue filter image. Each filter is stretched individually to maximize contrast.

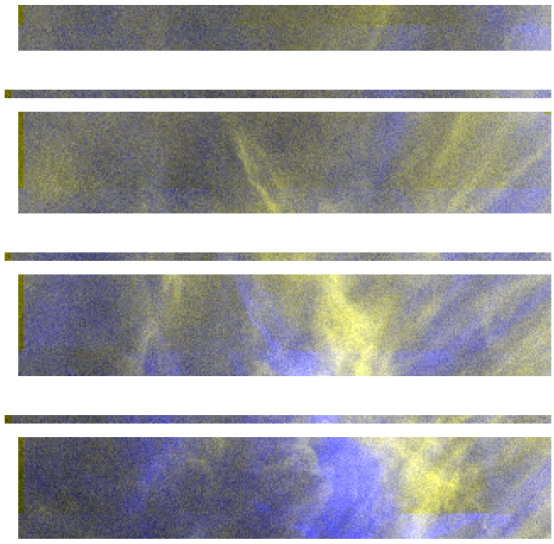

Figure 3.5: Co-alignment (*i.e.*, lack thereof) of the green and blue filters at the altitude of the local surface for the V04573003c ROI. Otherwise the same as Figure 3.4

details of RDR layout and radiance calibration.) The correlation-ROI is defined as a rectangular region within the green filter plane of the RDR. Normally, this region encompasses multiple framelets, and so when the correlation-ROI is displayed without map projection there are discontinuities at the framelet boundaries, as seen in Fig. 3.4.

To calculate the correlation, we first classify each RDR pixel as either valid or "null" (invalid). Valid pixels are used for calculating the correlation, and null pixels are excluded. The valid region of the framelets is defined so as to exclude the regions which commonly contain stray light artifacts (which are discussed in chapter 4). In framelet coordinates, the valid region is samples 160 – 840, lines 0 – 176 for 1x1 binning (divide these coordinates by 2 or 4 for 2x2 or 4x4 binning, respectively). Framelet line-sample coordinates are the same as the line and sample numbers in the RDR, except for a translation setting the lowest line number in a

framelet equal to 0. All pixels outside the valid region in a framelet are set to null. Pixels that are null in the original RDR are also classified as null.

Next, we project the blue filter data onto the green filter correlation-ROI with the given assumed altitude and velocity. To do so, we handle each green filter framelet within the correlation-ROI separately, and perform the following steps:

1. Calculate the latitude and longitude of each pixel in all blue filter framelets near to the green filter correlation-ROI. We do so following the method of the Unites States Geological Survey's ISIS software package [Torson and Becker, 1997; Gaddis et al., 1997, http://isis.astrogeology.usgs.gov/], which we have adapted to apply to targets of arbitrary altitude. This method consists of:

   (a) use the NASA Navigation and Ancillary Information Facility's SPICE data kernels and toolkit [Acton, 1996] to determine the exact position and attitude of ODY with respect to Mars, and from that the boresight vector of the THEMIS-VIS instrument;

   (b) use the ISIS package's THEMIS-VIS camera model, which includes a correction for optical distortion, to derive the pointing vector for each pixel from the boresight vector;

   (c) identify latitude and longitude of each pixel as the intersection of its pointing vector with an ellipsoid parallel to the Mars reference ellipsoid and at a given altitude above it.

   To verify this latitude and longitude calculation method, we have used it to map project THEMIS-VIS framelets from all filters onto the local surface elevation. These map projection tests show that we obtain results identical to the ISIS software. They also show occasional misalignment of surface

features by up to two 1x1 binning pixels. This means that inaccuracies in the ISIS camera model, or, less likely, the SPICE data kernels, may cause systematic errors in our altitude and velocity measurements of up to 1 km in altitude or 10 m/s in velocity. These potential errors are small compared to other uncertainties, and so we have not pursued this issue further.

2. Adjust the longitude of each blue filter pixel to compensate for the assumed zonal velocity. The necessary translation for the given velocity depends on the time elapsed between the green filter framelet and each blue filter framelet, and the result of the translation is that locations in the blue filter framelet are labeled with longitudes projected forward to the time that the green filter framelet was exposed.

3. Convert the latitude and velocity compensated longitude for each blue filter framelet pixel into a line and sample number in the reference frame of the green filter framelet's framelet coordinates. Our method in this step is also based on the ISIS software package:

   (a) calculate the vector from ODY to a pixel's latitude and compensated longitude on the ellipsoid at the given altitude;

   (b) apply the inverse of the ISIS package's THEMIS-VIS camera model to determine the sample and line number referenced to the green filter framelet's origin.

4. Interpolate the blue filter framelet image from the non-integer line and sample coordinates calculated for the blue filter pixels in the preceding step to the integer line-sample coordinates of the green filter framelet. We perform a separate linear interpolation for each blue filter framelet onto the green filter

framelet. Interpolated values are only created in the interior of each region covered by valid blue filter framelet pixels. Any pixels of the interpolated blue filter image which are interior to none of these regions are classified as null.

The above procedure results in a blue filter image that has been projected onto the green filter image for each framelet in the correlation-ROI. We then assemble these projected blue filter framelets into the blue filter projected image of the correlation-ROI. Any pixel which is null in either the projected blue filter correlation-ROI or the original green filter correlation-ROI is now set to null in both. Figures 3.4 and 3.5 are examples of the green filter correlation-ROI overlain with the projected blue filter correlation-ROI. Null pixels are blank (white) in these figures.

Electron counting noise in the image pixels leads to noise at a characterstic frequency in maps of correlation versus altitude and velocity. To address this, we apply a gaussian smoothing function, with a FWHM of 3 pixels in both the line and sample direction, to both the green filter and projected blue filter correlation-ROI images prior to calculating the correlation. We observe, *e.g.*, Figure 3.2, that this smoothing eliminates the noise in the correlation map and improves the overall precision of the measurement.

The altitude and velocity for the correlation-ROI are those that give the highest correlation between the projected blue filter and the green filter correlation-ROI. To find the maximum correlation, we first apply a downhill simplex non-linear optimization routine based on the "amoeba" routine from Press et al. [1992]. Our modified amoeba search halts when the change in altitude and velocity with subsequent iterations is less 0.5 km in altitude and 5 m/s in velocity. We then perform

an adaptive grid search on the altitude-velocity space surrounding the best fit produced by the amoeba search. The adaptive grid search allows us to map the probability density function of the altitude and velocity fit and acts as a consistency check for the amoeba search results. Our reported "best-fit" altitude and velocity estimate is the grid point with maximum correlation.

For the grid search, all grid points on the boundary of the grid must have a correlation more than three standard errors below that of the grid point with the best correlation. If this criterion is not satisfied for a particular boundary, the grid is expanded in that direction until it is satisfied. This adaptive grid search is performed up to three times with the grid sampling intervals decreased by a factor of two after each iteration. The initial grid sampling intervals are 2 km in altitude and 20 m/s in velocity. If the size of the adapted grid at the end of an iteration is larger than 10 grid points in either dimension, the grid search is halted in order to conserve computational resources.

One problem inherent in this correlation search method, as in all real-world localization problems, is that the movement of significant image features into or out of the region of interest as the altitude or velocity is changed leads to spurious changes in the correlation. For example, surface features are often visible beneath the high altitude cloud features that we are correlating. As the altitude and velocity are adjusted, a surface could move out of the region of valid pixels, artificially improving the calculated correlation. We mitigate this problem in two ways.

1. When performing the grid search, we require the correlation calculation at every grid point in an iteration to use the same set of correlation-ROI pixels. Effectively, we set all pixels that are null at any grid point to be null at all grid points. For the wider range of velocities and altitudes considered by the

amoeba search, this strict pixel validity requirement is infeasible, because it would lead to the majority of pixels being labeled null. The amoeba search therefore begins to use the strict pixel validity requirement only after it has narrowed down the altitude and velocity range. Once the amoeba search is altering the altitude and velocity estimates by less than 4 km and 40 m/s per iteration, it switches to the strict pixel validity requirement. For the amoeba search the validity requirement is cumulative once the strict requirement is invoked. Pixels that were invalid in one correlation calculation remain invalid for all future correlation calculations.

2. We project, as previously discussed, the blue filter framelets onto the green filter, instead of the the other way around. The above strict pixel validity requirement prevents any change in the features of the unprojected filter that are used in the correlation. However, it does not prevent features in the projected filter from moving into and out of the correlation-ROI. Surface features visible though the high-altitude cloud are much more prominent in the green filter than the blue filter. Thus, by using the blue filter as the projected filter, we minimize the significance of surface features causing spurious changes in the correlation by moving into or out of the correlation-ROI.

To calculate the standard error of the correlation coefficients at each grid point, we apply the bootstrap method [Efron and Tibshirani, 1993] at the grid point with the highest correlation. This bootstrap calculation generates 1000 simulated data sets by randomly sampling, with replacement, from the valid points in the correlation-ROI. The number of points in each simulated data set is equal to the number of points in the correlation-ROI. It then calculates the correlation for each

simulated data set, generating a set of correlation measurements that simulates repeating the correlation measurement 1000 times. The bootstrap standard error is the standard deviation of this simulated correlation data set. To identify the confidence intervals for the altitude and velocity estimates, we construct a test of the hypothesis that the correlation at a given grid point is not less than the maximum correlation, using the bootstrap standard error as the standard error for both the maximum correlation and the grid points. Our reported confidence intervals contain all points for which the probability of this hypothesis is greater than 5%. In Figure 3.2, the boundaries of this confidence interval are illustrated by a bold black line.

3.2.2 Altitude-velocity measurement: candidate selection

We have found that the most efficient and reliable method of identifying candidate mesospheric clouds in the THEMIS-VIS data set is to visually inspect each multi-spectral THEMIS-VIS image. Important obstacles to doing this automatically are 1) false positives caused by calibration artifacts and/or very low contrast image features; and 2) false negatives caused by the proximity of surface features to mesospheric cloud features.

We perform the search for candidates by map-projecting all multi-spectral THEMIS-VIS images onto the local surface and assembling the filters into enhanced color renderings such as those in Fig. 3.6. The enhanced color rendering simply assigns the blue, green, and "red" (654 nm) filter images to the blue, green and red channels of the standard RGB color space, stretching each filter image individually so that $\pm\ \alpha$ standard deviations from its mean fills the full dynamic range of the corresponding channel. If the "red" (654 nm) filter is not present in

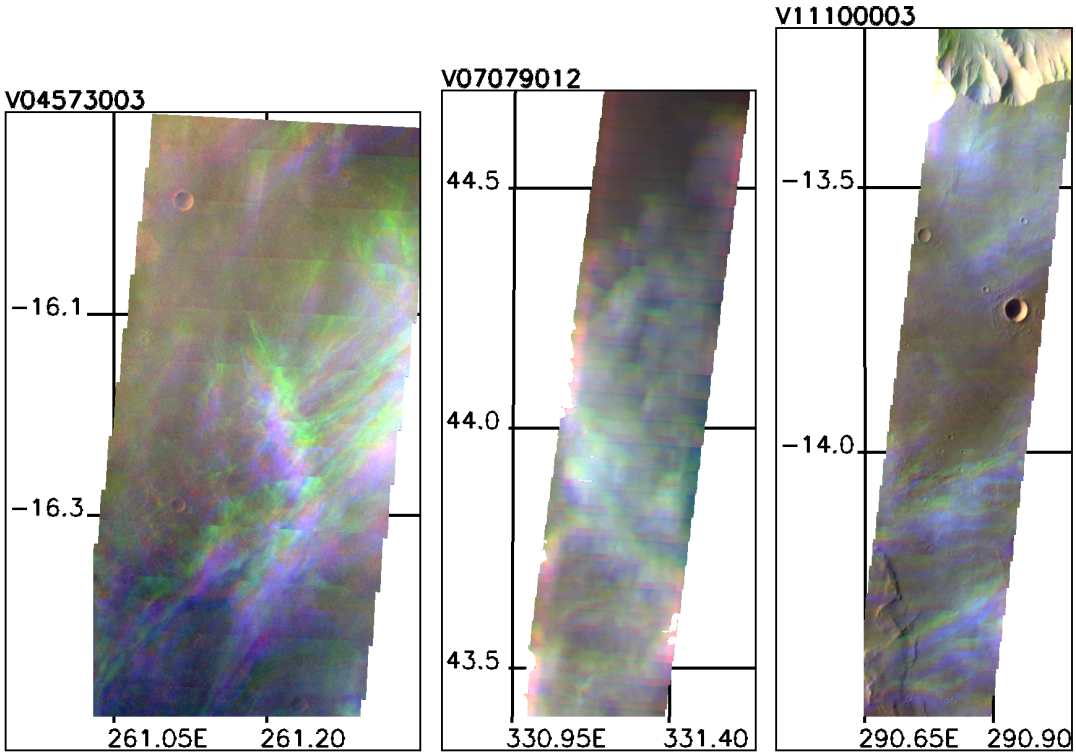

Figure 3.6: Example discovery images, showing the appearance of mesospheric clouds when map projected using the altitude of the local surface. These images use enhanced color rendering defined in the text, the grid lines give degrees east longitude and degrees north latitude, and the projection is sinusoidal.

a given image sequence, the 749 nm filter is used in its place. We chose α on a case by case basis, but typically use $\alpha = 2$. Any apparent motion relative to the surface shows up as a misalignment between the color channels. We rely on this misalignment to identify the mesospheric cloud candidates.

THEMIS-VIS sequence V04573003 (see Figure 3.6) was the first misalignment identified in the data set. Prior to its discovery, there was no expectation that mesospheric clouds would be observable by THEMIS-VIS. To date we have examined all multi-spectral THEMIS-VIS image sequences up to and including sequence V18086011 (acquired Jan. 11, 2006). Whenever an apparent color misalignment is

Table 3.1: Correlation-ROIs: equatorial mesospheric cloud candidates

ID[1]	L_s	N. lat[2]	E. lon[2]	° Inc.[2]	samples[3]	lines[3]
V04573003a	113.7	-16.661	261.136	80.2	59 – 421	1164 – 1398
V04573003b	113.7	-16.410	261.172	80.1	94 – 433	687 – 936
V04573003c	113.7	-16.240	261.196	80.0	84 – 420	322 – 655
V10324001a	18.3	-5.003	356.963	67.9	117 – 440	1263 – 1644
V10526009a	26.1	0.523	261.881	68.9	86 – 429	713 – 1126
V11100003a	47.5	-14.359	290.794	78.0	76 – 217	4685 – 4894
V11100003b	47.5	-14.071	290.818	77.9	44 – 203	4393 – 4639
V11100003c	47.5	-13.562	290.891	77.8	53 – 220	3918 – 4151
V12922001a	113.8	-6.201	292.486	83.0	101 – 334	268 – 658
V13072001a	119.6	-5.308	285.542	82.9	80 – 257	1194 – 1333
V13072001b	119.6	-5.208	285.576	82.9	67 – 341	802 – 1354

[1]THEMIS-VIS sequence number, with a lower case letter appended to distinguish individual ROIs

[2]At the center of the ROI

[3]Line/sample numbers as stored in RDR; first sample, line labeled 0, 0.

identified, we then select one or more correlation-ROIs from the RDR of that imaging sequence. The rectangular correlation-ROIs are selected to enclose a region of morphologically similar misaligned features while avoiding, as much as possible, high contrast surface features. When the misaligned features cover a large portion of the sequence, or when there are multiple morphologically distinct groups of apparent high altitude clouds, we select multiple correlation-ROIs in that sequence. We have also selected a few cloud features with no apparent misalignment to serve as a control for our identification method. Tables 3.1 – 3.4 list all of the candidate correlation-ROIs, including the controls.

The selected candidate correlation-ROIs range from cases of obvious cloud features with dramatic misalignment, to cases with barely visible features and questionable misalignment. Even though we take care to include marginal cases as candidates, this detection method is obviously biased towards the highest altitude clouds. This is one reason why we have chosen to limit this work to mesospheric clouds. Even though our measurement technique is precise enough to detect the

Table 3.2: Correlation-ROIs: mesospheric cloud candidates L_s 200° – 250°

ID	L_s	N. lat	E. lon	° Inc.	samples	lines
V06705017a	206.1	49.723	318.187	94.9	41 – 217	1067 – 1621
V06717019a	206.7	51.174	332.432	95.4	22 – 226	67 – 418
V06768014a	209.3	44.059	300.252	94.1	30 – 223	536 – 894
V06768014b	209.3	44.369	300.311	94.2	29 – 230	311 – 523
V06768014c	209.3	44.622	300.359	94.2	26 – 232	51 – 299
V06793016a	210.5	40.500	298.631	93.3	41 – 228	2296 – 2891
V06793016b	210.5	42.161	298.929	93.8	25 – 224	691 – 1305
V06905017a	216.2	45.794	309.566	96.1	24 – 224	3000 – 3338
V06905017b	216.2	46.597	309.723	96.4	26 – 224	2216 – 2574
V06917014a	216.8	44.586	323.246	95.8	47 – 211	678 – 935
V06930045a	217.4	44.660	308.360	96.0	40 – 218	4148 – 4599
V06930045b	217.4	45.830	308.585	96.4	42 – 221	2921 – 3574
V06930045c	217.4	47.637	308.944	97.0	36 – 222	1354 – 1666
V07005019a	221.3	43.785	305.275	96.4	31 – 230	527 – 974
V07005019b	221.3	44.324	305.377	96.6	35 – 222	34 – 424
V07029045a	222.5	46.433	333.637	97.5	38 – 215	3040 – 3438
V07029045b	222.5	48.225	333.997	98.2	36 – 214	1380 – 1650
V07067015a	224.4	44.425	317.417	97.1	41 – 213	2616 – 2867
V07079012a	225.0	42.557	330.996	96.5	32 – 209	4287 – 4671
V07079012b	225.0	42.986	331.075	96.7	31 – 217	3865 – 4259
V07079012c	225.0	43.848	331.234	97.0	34 – 223	3042 – 3430
V07079012d	225.0	44.416	331.342	97.2	33 – 224	2433 – 2942
V07080012a	225.1	42.730	302.198	96.6	39 – 212	2671 – 2969
V07080012b	225.1	44.117	302.479	97.1	36 – 219	1273 – 1704
V07105019a	226.4	41.141	300.979	96.2	50 – 209	3150 – 3376
V07105019b	226.4	41.700	301.072	96.4	48 – 201	2553 – 2900
V07166022a	229.5	39.663	341.549	95.9	47 – 202	3779 – 4009
V07166022b	229.5	40.575	341.708	96.3	48 – 202	2891 – 3144
V07166022c	229.5	41.855	341.928	96.8	54 – 194	1638 – 1939
V07166022d	229.5	43.457	342.231	97.4	41 – 214	79 – 411
V07467019a	245.2	48.212	302.699	100.5	39 – 213	730 – 1138
V07504020a	247.1	46.357	315.269	99.7	44 – 208	1160 – 1385

Table 3.3: Correlation-ROIs: mesospheric cloud candidates L_s 250° – 300°

ID	L_s	N. lat	E. lon	° Inc.	samples	lines
V07804021a	262.8	40.115	302.462	95.6	36 – 206	2573 – 2901
V07804021b	262.8	38.696	302.218	94.9	34 – 220	3856 – 4343
V07829020a	264.1	47.187	302.854	99.3	43 – 206	3011 – 3608
V07829020b	264.1	48.250	303.066	99.8	41 – 210	2166 – 2411
V07890023a	267.2	41.708	342.524	95.8	33 – 212	2326 – 2701
V07941020a	269.9	44.094	312.117	96.7	37 – 216	2662 – 2989
V07941020b	269.9	45.354	312.350	97.4	38 – 218	1462 – 1764
V07954018a	270.6	40.326	296.543	94.5	40 – 214	2986 – 3474
V08028019a	274.4	43.697	322.877	95.6	47 – 203	3090 – 3563
V08028019b	274.4	44.219	322.975	95.9	43 – 208	2556 – 3090
V08057021a	275.9	38.986	205.726	92.7	40 – 209	529 – 855
V08140022a	280.1	42.678	332.560	93.9	46 – 202	2123 – 2484
V08140022b	280.1	44.092	332.820	94.7	43 – 211	695 – 1190
V08141026a	280.2	41.550	303.553	93.2	18 – 232	1140 – 1418
V08141026b	280.2	42.297	303.684	93.7	51 – 221	437 – 685
V08165025a	281.4	44.069	331.810	94.4	42 – 208	830 – 1528
V08165025b	281.4	45.069	331.999	95.0	43 – 211	69 – 362
V08166020a	281.5	42.530	302.725	93.5	38 – 207	215 – 568
V08178017a	282.1	53.987	318.934	99.9	38 – 216	683 – 989
V08266019a	286.5	45.988	299.321	94.3	42 – 212	2503 – 2769
V08266019b	286.5	46.449	299.402	94.6	44 – 202	2020 – 2366
V08266019c	286.5	46.922	299.494	94.9	40 – 204	1595 – 1881
V08266019d	286.5	47.454	299.604	95.2	36 – 217	934 – 1516
V08266019e	286.5	48.116	299.733	95.5	36 – 209	372 – 800
V08278020a	287.2	48.334	313.667	95.5	23 – 225	1168 – 1508
V08278020b	287.2	49.130	313.828	96.0	23 – 221	351 – 791
V08290017a	287.8	44.911	326.900	93.4	40 – 212	492 – 992
V08302017a	288.4	46.826	341.161	94.3	46 – 210	645 – 1131
V08302017b	288.4	47.426	341.275	94.7	48 – 207	28 – 591
V08303022a	288.4	49.074	312.783	95.6	24 – 226	2800 – 3328
V08303022b	288.4	51.293	313.244	96.9	28 – 221	576 – 1276
V08428013a	294.7	50.802	308.871	95.0	33 – 215	4101 – 5054
V08428013b	294.7	52.979	309.343	96.3	36 – 218	2176 – 2773
V08503016a	298.4	48.700	305.923	92.8	49 – 211	2935 – 3375
V08503016b	298.4	49.495	306.085	93.3	50 – 212	2119 – 2658

Table 3.4: Correlation-ROIs: controls

ID	L_s	N. lat	E. lon	° Inc.	samples	lines
V08806011a	313.1	29.804	203.173	78.0	41 – 213	2628 – 3104
V09543022a	346.5	32.091	182.502	71.7	29 – 213	1069 – 1465
V09550015a	346.8	43.752	342.489	76.5	39 – 211	1247 – 1509
V09575016a	347.8	48.573	342.225	78.3	37 – 220	852 – 1189
V09912021a	1.9	52.484	341.718	76.2	40 – 211	809 – 1184

altitude of lower atmosphere clouds, the misalignment is too small to be noted by visual inspection. In fact, some of the control correlation-ROIs yield altitudes in the ∼10 km range (none exceeded 13 km). It may ultimately be feasible to define correlation-ROIs for all cloud features in the THEMIS-VIS data set, but reliably distinguishing low-altitude clouds from surface features will be problematic, especially when the signal is very low as it often is in winter mid-latitudes. Since such an effort would be expensive in terms of both human time and computer time, we have decided that it is beyond the scope of this work. Thus we restrict ourselves to 40+ km altitudes in order to ensure that we have a nearly complete census.

In order to decide whether our correlation algorithm has successfully identified a cloud's altitude and velocity we look at the width of the confidence interval for the altitude solution. In some cases, the algorithm will give a non-zero altitude-velocity solution for which visual inspection shows no improvement in the co-alignment of what were believed to be the cloud features. These algorithm failures appear to be caused by calibration artifacts in the THEMIS-VIS images (see Chapter 4), which will dominate scene cross-correlation if the cloud contrast is very low or nonexistent. However, we invariably find these apparent algorithm failures to have confidence intervals wider than 10 km in altitude, and so we chose at 10 km as our cutoff in order to eliminate false positives.

3.3 Altitude-velocity results

In Table 3.5 – 3.8 we show results from all correlation-ROIs that yield altitudes greater than 40 km and have confidence intervals less than 10 km wide. Each result is given an ID which consists of the THEMIS-VIS sequence number followed by a lower case letter specifying the ROI. All of the observed mesospheric clouds fall

Table 3.5: Altitude and velocity measurements: equatorial

ID	L_s	N. Lat	E. Lon	° Inc.	Measured				WRF GCM	
					Altitude *km*		Zonal Velocity *m/s*		Zonal Velocity *m/s*	
					Best fit	Conf. interval	Best fit	Conf. interval	Mean	$\pm 2\sigma$
V04573003a	113	−16	261	80.2	79.5	79.0 – 80.0	−15	−20 – −10	2	−18 – 22
V04573003b	113	−16	261	80.1	80.0	79.0 – 81.0	−5	−15 – 0	2	−18 – 22
V04573003c	113	−16	261	80.0	80.0	79.5 – 80.5	−5	−5 – 0	2	−18 – 22
V10526009a	26	0	261	68.9	74.0	71.5 – 76.0	−90	−110 – −70	0	−19 – 18
V11100003a	47	−14	290	78.0	61.0	59.5 – 62.0	−65	−80 – −50	10	−4 – 24
V11100003b	47	−14	290	77.9	61.5	61.0 – 62.5	−50	−60 – −35	10	−4 – 24
V11100003c	47	−13	290	77.8	62.0	61.0 – 63.5	−65	−80 – −45	10	−4 – 24
V12922001a	113	−6	292	83.0	57.0	56.5 – 57.5	−45	−50 – −40	31	9 – 53
V13072001a	119	−5	285	82.9	60.0	58.5 – 61.0	−45	−60 – −30	0	−26 – 27
V13072001b	119	−5	285	82.9	59.0	57.5 – 60.5	−70	−85 – −55	13	−13 – 39

into one of two distinct classes: equatorial clouds observed in the solar longitude (L_s) range 0° – 180°, listed in Table 3.5, and northern mid-latitude clouds observed in L_s 200° – 300°, listed in Tables 3.6 – 3.8. The northern mid-latitude clouds are organized into sub-classes by our assessment of their morphology. Table 3.6 shows the "clumpy" subclass, which consists of any cloud that shows extensive small scale structure but no obvious larger scale organization. In Table 3.7 we list "linear" clouds, those that are organized into one or more southwest-to-northeast trending linear features. The final northern mid-latitude subclass, listed in Table 3.8, consists of "linear periodic" clouds, those whose multiple linear features show regular spacing. The equatorial mesospheric clouds also show some morphological diversity. Three of them, V04573003, V10526009, and V11100003 show long narrow filamentary structures, while the other two do not. Example map projected images from both classes and each subclass are shown in Figs. 3.7 – 3.10.

3.3.1 Comparisons with an atmospheric global circulation model (GCM)

In order to put our equatorial velocity measurements into context, we compare them with an atmospheric general circulation model (GCM). We have chosen the

Table 3.6: Altitude and velocity measurements: mid-latitude, "clumpy" class

ID	L_s	N. Lat	E. Lon	° Inc.	Measured				WRF GCM	
					Altitude *km*		Zonal Velocity *m/s*		Zonal Velocity *m/s*	
					Best fit	Conf. interval	Best fit	Conf. interval	Mean	$\pm 2\sigma$
V06768014a	209	44	300	94.1	52.5	50.5 – 54.5	85	65 – 100	75	49 – 101
V06793016b	210	42	298	93.8	48.5	47.0 – 49.5	75	65 – 90	55	35 – 76
V06905017a	216	45	309	96.1	62.0	60.0 – 63.5	60	45 – 80	57	21 – 92
V06905017b	216	46	309	96.4	61.0	59.5 – 62.0	65	50 – 80	57	21 – 92
V06930045b	217	45	308	96.4	70.0	68.5 – 72.0	65	50 – 75	51	27 – 74
V06930045c	217	47	308	97.0	73.5	71.5 – 75.0	75	60 – 90	88	25 – 150
V07005019a	221	43	305	96.4	64.0	62.5 – 65.5	25	10 – 35	47	9 – 86
V07005019b	221	44	305	96.6	65.0	63.5 – 67.0	35	25 – 50	47	9 – 86
V07029045a	222	46	333	97.5	63.0	60.5 – 65.5	40	20 – 55	60	31 – 89
V07079012a	225	42	330	96.5	70.5	69.0 – 72.0	0	−10 – 10	40	−7 – 87
V07079012b	225	42	331	96.7	69.5	68.5 – 70.5	0	−10 – 5	46	−5 – 97
V07079012c	225	43	331	97.0	68.0	67.0 – 69.0	−10	−20 – −5	46	−5 – 97
V07079012d	225	44	331	97.2	69.5	68.5 – 70.5	25	20 – 30	46	−5 – 97
V07105019b	226	41	301	96.4	70.0	65.0 – 74.0	−30	−70 – 5	38	5 – 71
V07166022c	229	41	341	96.8	60.0	55.0 – 64.0	50	20 – 80	28	−20 – 75
V07166022d	229	43	342	97.4	59.0	57.0 – 61.5	55	35 – 70	53	9 – 97
V07804021a	262	40	302	95.6	51.5	51.0 – 52.0	30	25 – 35	43	0 – 85
V07804021b	262	38	302	94.9	54.0	53.0 – 54.5	15	10 – 25	40	1 – 79
V07941020a	269	44	312	96.7	49.0	47.5 – 50.0	35	25 – 45	68	35 – 101
V07941020b	269	45	312	97.4	49.5	47.0 – 51.5	30	15 – 45	80	50 – 110
V07954018a	270	40	296	94.5	52.5	51.0 – 53.0	0	−10 – 10	45	16 – 75
V08140022a	280	42	332	93.9	49.5	48.0 – 51.5	60	45 – 75	89	57 – 121
V08140022b	280	44	332	94.7	48.0	46.0 – 50.0	60	50 – 75	89	57 – 121
V08266019a	286	45	299	94.3	49.5	46.5 – 52.0	50	30 – 75	77	43 – 111
V08266019b	286	46	299	94.6	50.5	49.0 – 52.0	35	20 – 50	77	43 – 111
V08266019c	286	46	299	94.9	51.5	49.5 – 53.0	40	20 – 60	85	58 – 111
V08266019d	286	47	299	95.2	47.0	45.0 – 49.0	45	25 – 60	88	61 – 115
V08278020a	287	48	313	95.5	47.0	45.5 – 48.5	120	105 – 130	99	59 – 139
V08278020b	287	49	313	96.0	47.0	46.0 – 48.5	120	105 – 130	109	65 – 154

Table 3.7: Altitude and velocity measurements: mid-latitude, "linear" class

ID	L_s	N. Lat	E. Lon	° Inc.	Measured				WRF GCM	
					Altitude *km*		Zonal Velocity *m/s*		Zonal Velocity *m/s*	
					Best fit	Conf. interval	Best fit	Conf. interval	Mean	$\pm 2\sigma$
V06705017a	206	49	318	94.9	50.0	48.0 – 52.0	85	60 – 105	105	82 – 128
V06793016a	210	40	298	93.3	47.5	46.0 – 49.5	55	35 – 75	55	35 – 75
V07467019a	245	48	302	100.5	44.5	43.5 – 46.0	95	80 – 105	97	49 – 146
V07829020b	264	48	303	99.8	41.0	38.0 – 43.5	50	15 – 90	105	59 – 152
V08028019a	274	43	322	95.6	60.0	56.5 – 61.0	25	−10 – 55	54	16 – 93
V08028019b	274	44	322	95.9	50.5	47.5 – 51.5	50	25 – 75	75	38 – 112
V08057021a	275	38	205	92.7	46.5	43.5 – 51.0	35	10 – 75	46	9 – 82
V08302017a	288	46	341	94.3	50.5	47.5 – 54.0	55	20 – 80	88	62 – 113
V08302017b	288	47	341	94.7	46.5	43.5 – 48.0	85	65 – 110	95	70 – 121
V08303022a	288	49	312	95.6	46.0	45.0 – 47.5	110	100 – 120	101	61 – 141
V08503016b	298	49	306	93.3	45.0	44.5 – 46.0	60	45 – 70	106	71 – 141

Table 3.8: Altitude and velocity measurements: mid-latitude, "linear periodic" class

ID	L_s	N. Lat	E. Lon	° Inc.	Measured				WRF GCM	
					Altitude *km*		Zonal Velocity *m/s*		Zonal Velocity *m/s*	
					Best fit	Conf. interval	Best fit	Conf. interval	Mean	$\pm 2\sigma$
V07890023a	267	42	343	95.8	54.0	51.0 – 56.0	30	10 – 50	36	−14 – 87
V08141026a	280	42	304	93.2	51.5	48.5 – 53.0	70	50 – 85	60	41 – 79
V08141026b	280	42	304	93.7	52.5	50.5 – 54.5	55	40 – 70	60	41 – 79
V08165025b	281	45	332	95.0	49.0	45.0 – 53.0	50	10 – 90	88	55 – 122
V08290017a	288	45	327	93.4	48.0	44.5 – 52.0	95	60 – 125	85	51 – 119
V08303022b	288	51	313	96.9	56.5	53.5 – 59.0	60	30 – 85	97	55 – 138
V08503016a	298	49	306	92.8	44.5	44.0 – 45.0	45	25 – 70	99	62 – 136

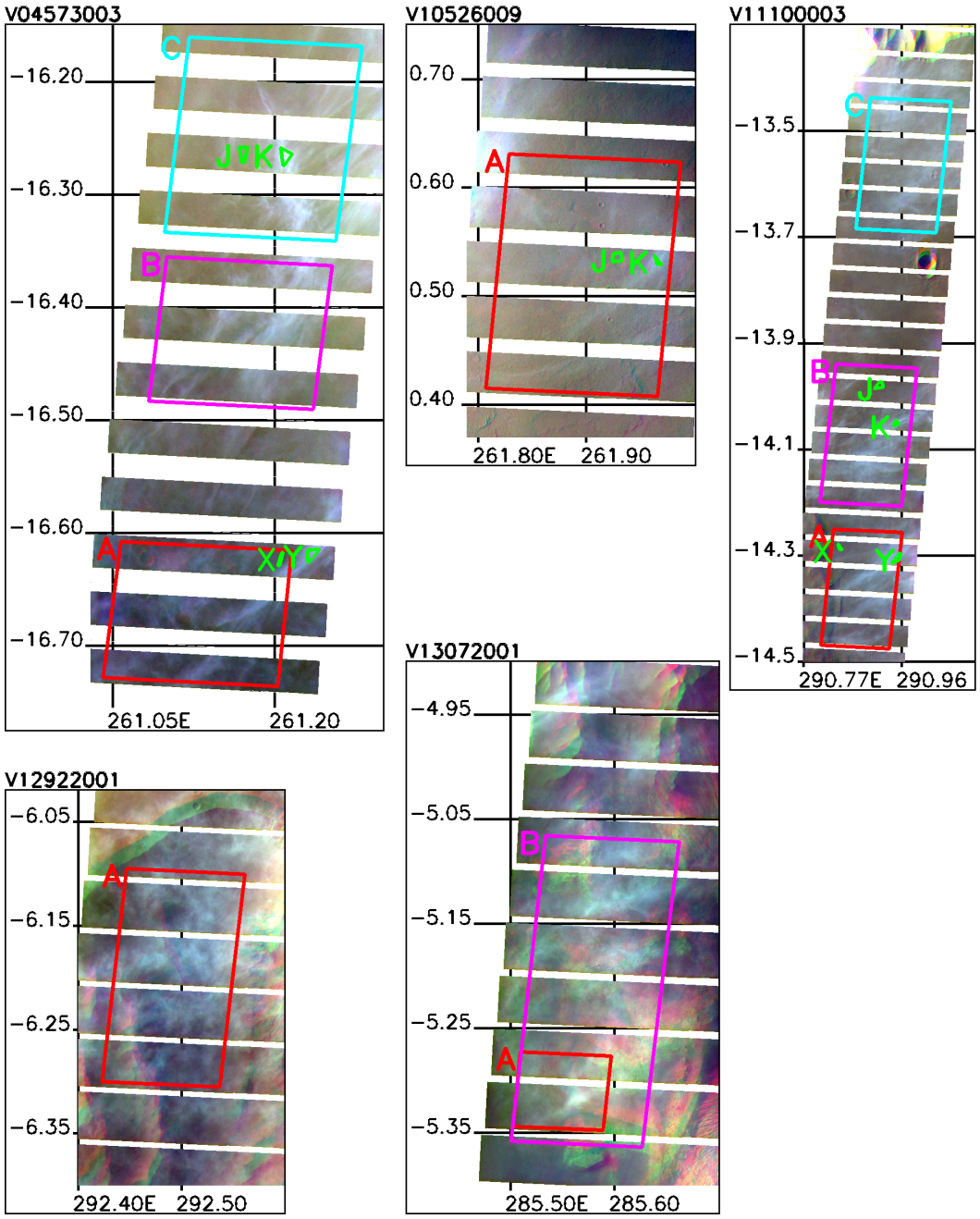

Figure 3.7: Examples of equatorial mesospheric clouds, using enhanced color as defined in the text. Each image is map projected (the projection is sinusoidal) using the best fit altitude and velocity. The grid lines give degrees east longitude and degrees north latitude. The large ROIs indicated by letters A – D are the correlation-ROIs, and their letters correspond to those in the first column of tables 3.1 – 3.8. The smaller ROIs are cloud / cloud-free radiance pairs. Their labels correspond to the columns in Tables 3.9 – 3.11.

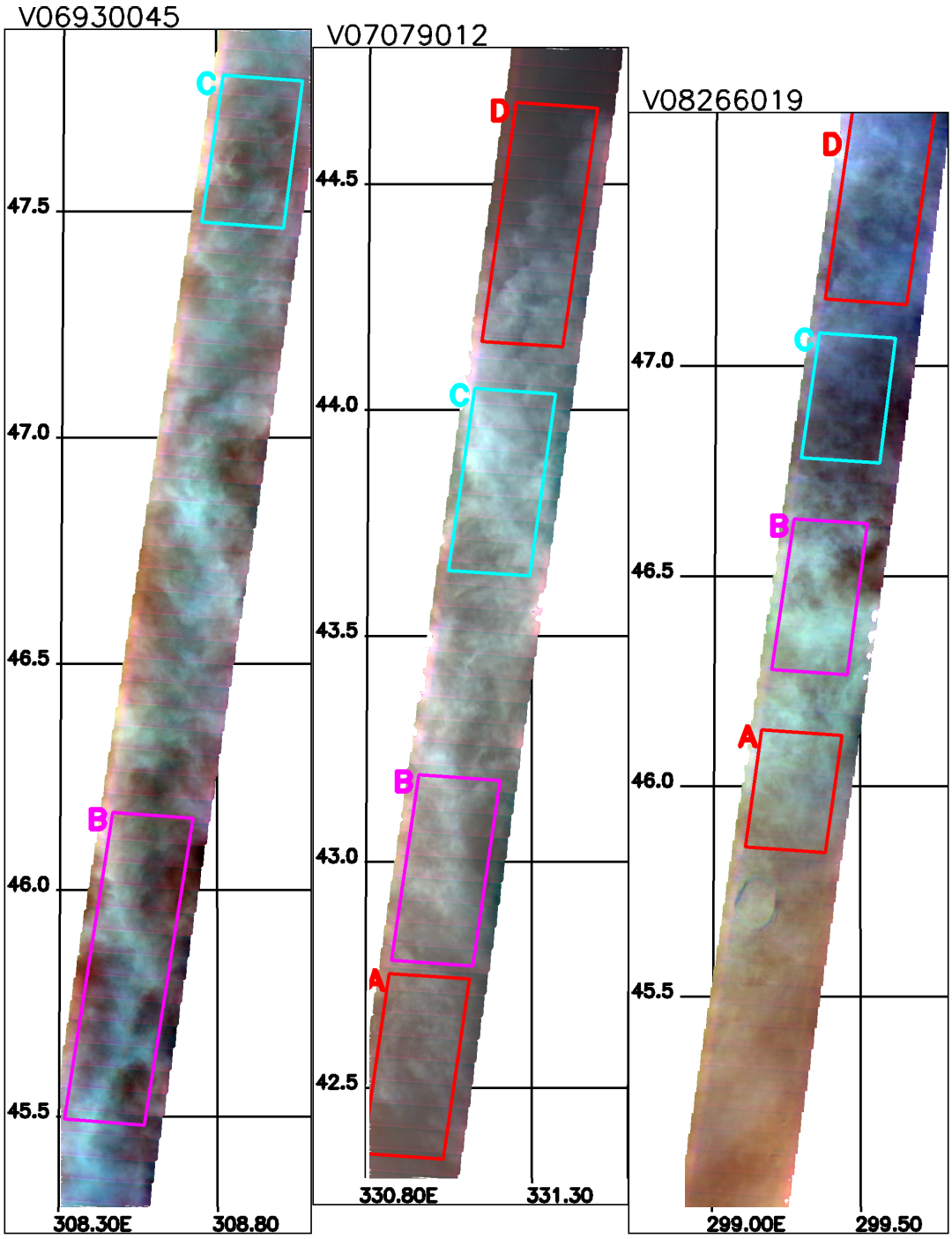

Figure 3.8: Examples of the "clumpy" class of mid-latitude mesospheric clouds. For V06930045 and V07079012 we have filled in narrow gaps between framelets by extrapolating from neighboring pixels. Otherwise the same as Figure 3.7.

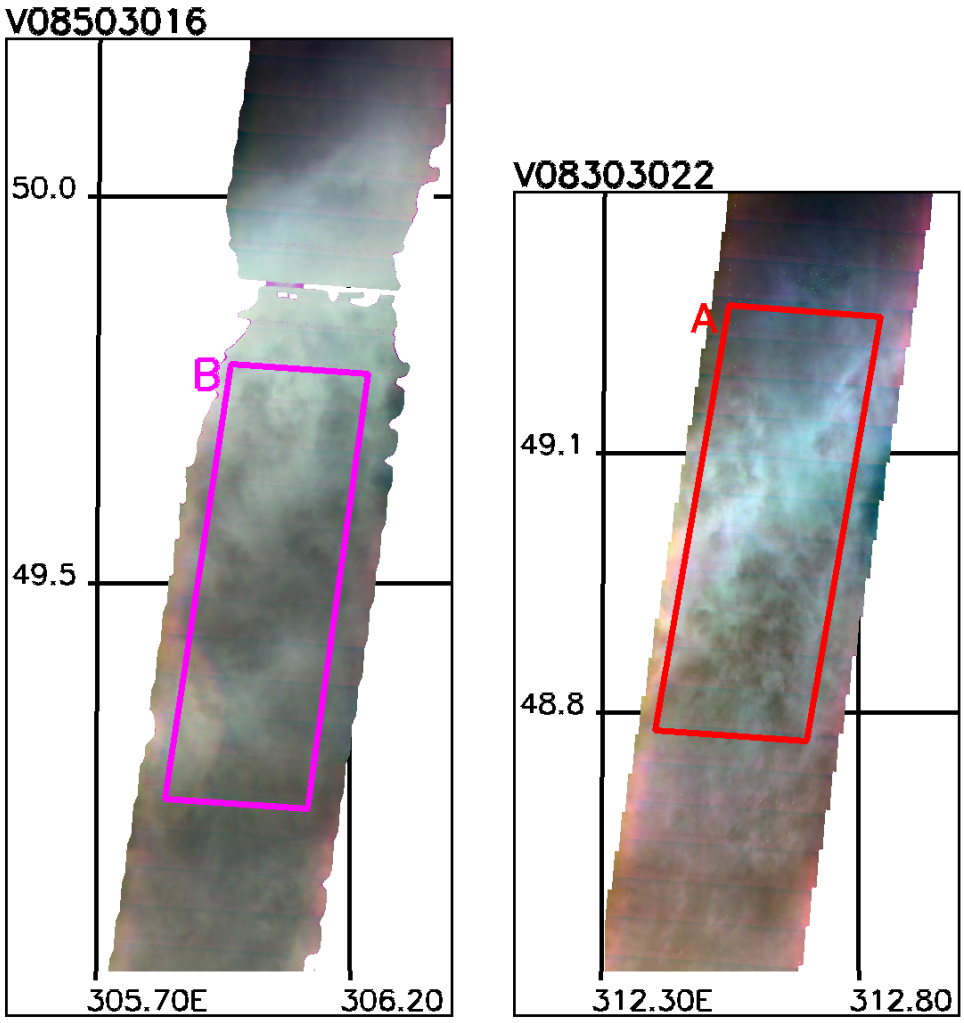

Figure 3.9: Examples of the "linear" class of mid-latitude mesospheric clouds. In V08503016, there are some gaps in the image caused by saturated pixels in the green filter image. We have extrapolated to fill the narrower gaps. Otherwise the same as Figure 3.7.

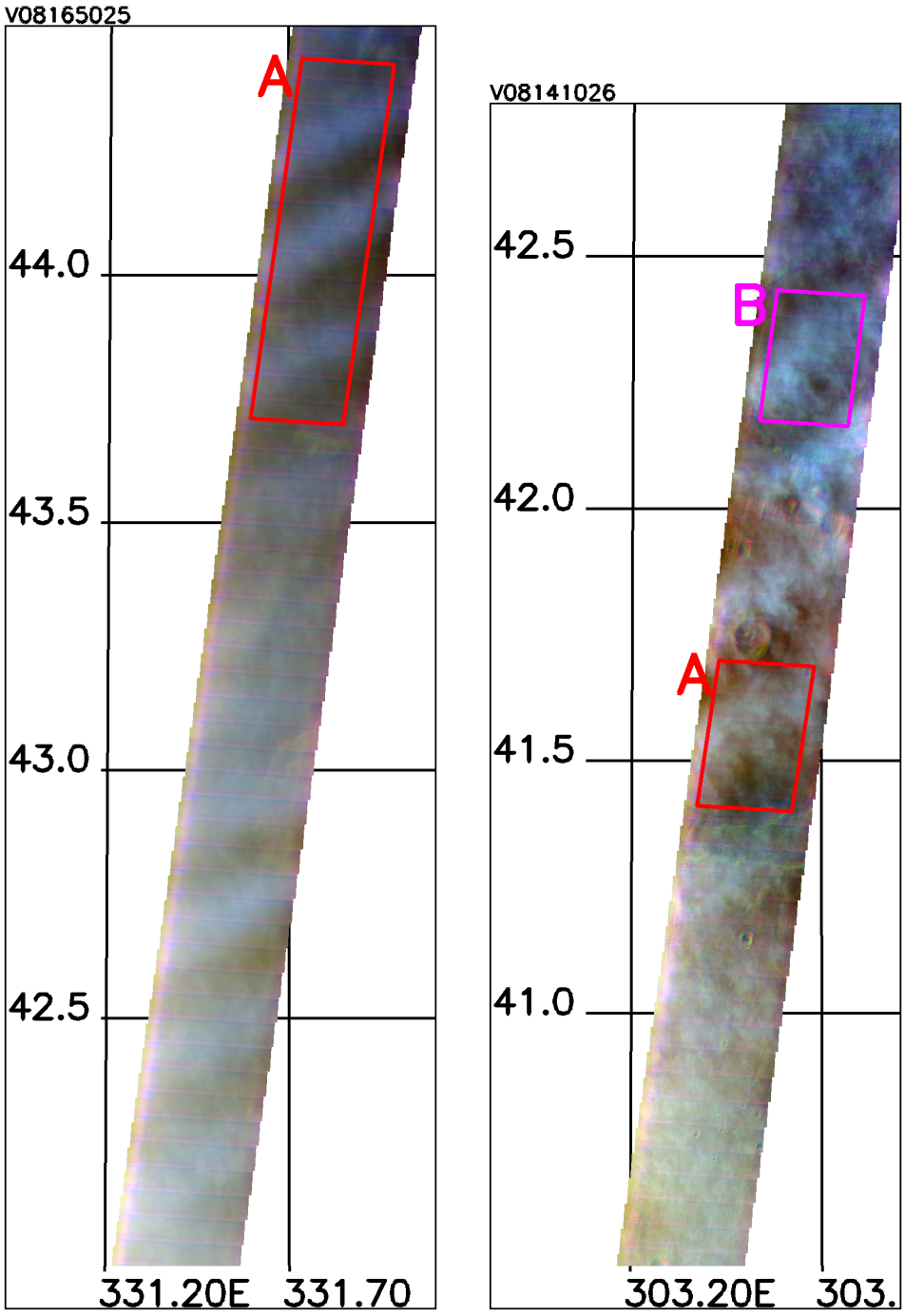

Figure 3.10: Examples of the "linear-periodic" class of mid-latitude mesospheric clouds. Otherwise the same as Figure 3.7.

Planetary WRF, described by Toigo et al. [2005]. The WRF model uses a terrain-following, pressure-based vertical coordinate, with the top layer of the model located at about 80 km above the reference geoid. All Mars GCMs include some form of damping near their model tops to parameterize gravity wave drag. The Toigo et al. [2005] WRF model uses Rayleigh drag, the most common form of simple momentum and temperature damping. Specifically, they spread the damping over the three uppermost model layers, with the shortest time constant (largest drag coefficient) in the topmost layer. The time constant grows by a factor of three for each of the two underlying layers. This form of gravity wave drag parameterization is not optimized for accuracy in the upper GCM layers; its main purpose is to prevent gravity waves from reflecting off the top of the model and influencing the lower layers that have historically been most readily compared with observations. Since direct measurements of winds are scarce at any altitude on Mars, it is difficult to match a GCM to the wind field. Our chosen model does a good job of representing the observed mean zonal temperature structure, and so it produces mean zonal flows consistent with the thermal winds that can be inferred from these temperature observations. Thus, our chosen model represents a reasonable point of comparison with our wind measurements, but we must keep in mind the limitations of the model, especially the lack of relevant wind data to validate itself against.

To make the wind velocity comparisons, we choose the final year of a 10 year model run and select the vertical level closest to the measured altitude of the given mesospheric cloud. (The closest model level is always within 5 km of the mesospheric cloud altitude.) We then select the four model grid points closest to the latitude-longitude coordinates of the mesospheric cloud from each of the two

local times nearest to the that of the cloud image, and from each of four days surrounding the the L_s of the cloud image. We use all of these points to generate the mean and standard deviation of the model winds shown in Tables 3.5 – 3.8.

3.3.2 Equatorial mesospheric clouds

The equatorial mesospheric clouds were discovered by pure chance in images targeted at surface features. Figure 3.11 shows the spatial distribution of these equatorial clouds, comparing it with the distribution of all THEMIS-VIS image sequences that "could have detected" mesospheric clouds had they been present. To detect mesospheric clouds, a THEMIS-VIS image sequence must obviously be multispectral. In addition, for purposes of Figure 3.11, we have also excluded types of image sequences which we believe are substantially less likely to detect mesospheric clouds. These are: 1) images that do not have both the green and blue filters, since these filters provide much better contrast for cloud features; 2) 1x1 binning sequences, since the spatial coverage of an individual 1x1 sequence is very small; 3) short exposure times ($<$ 1 msec for 2x2 binning or 0.5 msec for 4x4 binning), since such sequences have serious stray light problems that tend to obscure subtle features; 4) incidence angles $>$ 90 °, since these tend to have low signal levels.

Out of the 2048 could-have-detected sequences plotted in the lower panel of Figure 3.11, we have found 5 with equatorial mesospheric clouds — a detection rate of 0.24% (0.48% considering only images within 20 degrees of the equator). There appears to be an association with the eastern Tharsis high altitude terrain and with Valles Marineris. With only a slight overabundance of the could-have-detected sequences in the 260 - 295 longitude range of the detections, these five detections

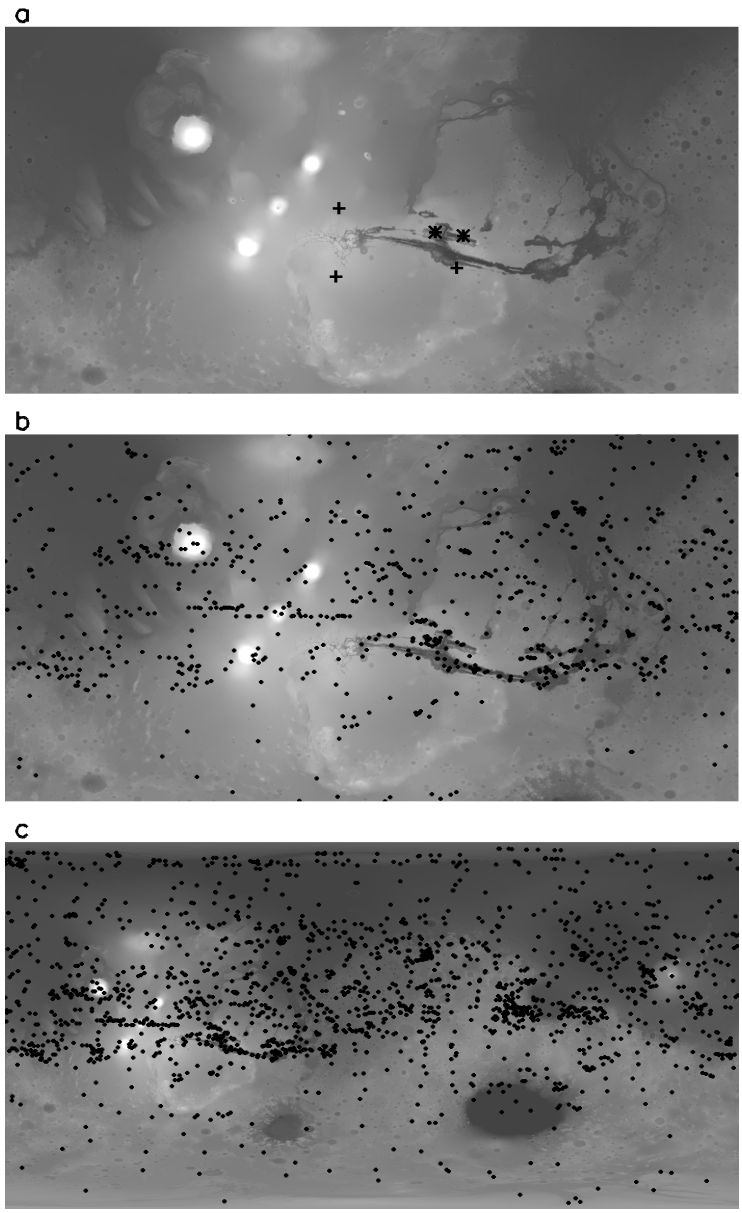

Figure 3.11: (a) Map of the locations where we have observed equatorial mesospheric clouds. Clouds in the "filamentary" subclass are marked with a +, and non-filamentary clouds are marked with an asterisk. In cases where there are multiple mesospheric cloud ROIs in a single THEMIS-VIS image sequence, we have plotted only one from each subclass. (b) and (c) Location of all of the THEMIS-VIS images that could have detected mesospheric clouds. (c) shows the entire planet. (b) shows the same region of the planet as in panel (a). The background for each map is grayscale-coded topography as measured by the Mars Orbiter Laser Altimeter.

represent a significant clustering. All of these detections fall within one of the longitude bands where Clancy et al. [2006] report detached hazes are prevalent. It is also intriguing that none of our equatorial mesospheric cloud detections fall in the L_s 60° – 100° season where both Jaquin [1989] and Clancy et al. [2006] note an absence of high altitude hazes, but given our small number of detections the absence of one in any particular period cannot be considered a significant result.

The morphology of the equatorial mesospheric clouds is reminiscent of the Pathfinder "discrete linear clouds" discussed by Clancy and Sandor [1998] (see their Figure 3), in that they are composed of "filaments" or "lineations." However, this is a highly subjective comparison, especially given the radically different viewing geometries of Pathfinder and THEMIS-VIS. Thus the THEMIS-VIS measurements cannot not confirm the hypothesis that the particular clouds observed by Pathfinder were at mesospheric altitudes. They do however confirm the more general hypothesis that mesospheric aerosols form not just detached hazes but also discrete, highly structured clouds.

Comparison with GCM

Figure 3.12 shows a comparison of the model winds and observed mesospheric cloud velocities for both the equatorial clouds and the mid-northern latitude clouds (further described in section 3.3.3). While the agreement is generally quite good for the mid-northern latitude data, the agreement is not as good with the sparser equatorial data. The GCM generally predicts a near absence of winds at these altitudes (and season and time of day). The lack of significant wind at these altitudes and location is also predicted by other Martian GCMs, and is not merely a product of the GCM we have chosen here. Four out of the five measurements indicate the

presence of an easterly zonal wind. We can think of at least two possible sources of this disagreement with this and other models. One possible reason for this disagreement is inaccuracies in the model, or, more specifically, the parameterization of gravity wave drag at high altitudes. As previously discussed, the lack of systematic direct wind measurements, at any altitude, make it difficult to improve models in this direction. The measurements presented here will hopefully help with progress in this area. Another possible cause of the disagreement is that the equatorial wind measurements come from a very localized region, with dramatic changes in topography over relatively short distances. Considering the inherently low horizontal resolution of a typical GCM, the the model may not be capturing higher-resolution wind flows being generated by topography and/or eddies.

3.3.3 Mid-latitude mesospheric clouds

All of the mid-latitude mesospheric clouds were discovered serendipitously in images targeted as part of a "Frontal Storm Survey" [Richardson et al., 2006] which had the intended goal of capturing high-resolution views of the major storm systems that are occasionally observed by MOC-WA [Wang and Ingersoll, 2002; Wang et al., 2005] in northern mid-latitudes in the fall and winter. THEMIS-VIS did not succeed in capturing any usable images of the intended storm systems, due in part perhaps to the late afternoon solar time of the THEMIS-VIS observations. (THEMIS has to date been constrained to a nadir-pointing configuration, and the ODY orbit is sun-synchronous with a local solar time allowed to drift between 4 and 5:30 pm.) This late local time meant that all THEMIS-VIS images acquired of the two targeted storm track regions in Acidalia and Amazonis Planitiae between L_s 200° and L_s 300° were of surfaces with incidence angles greater than 90°. All

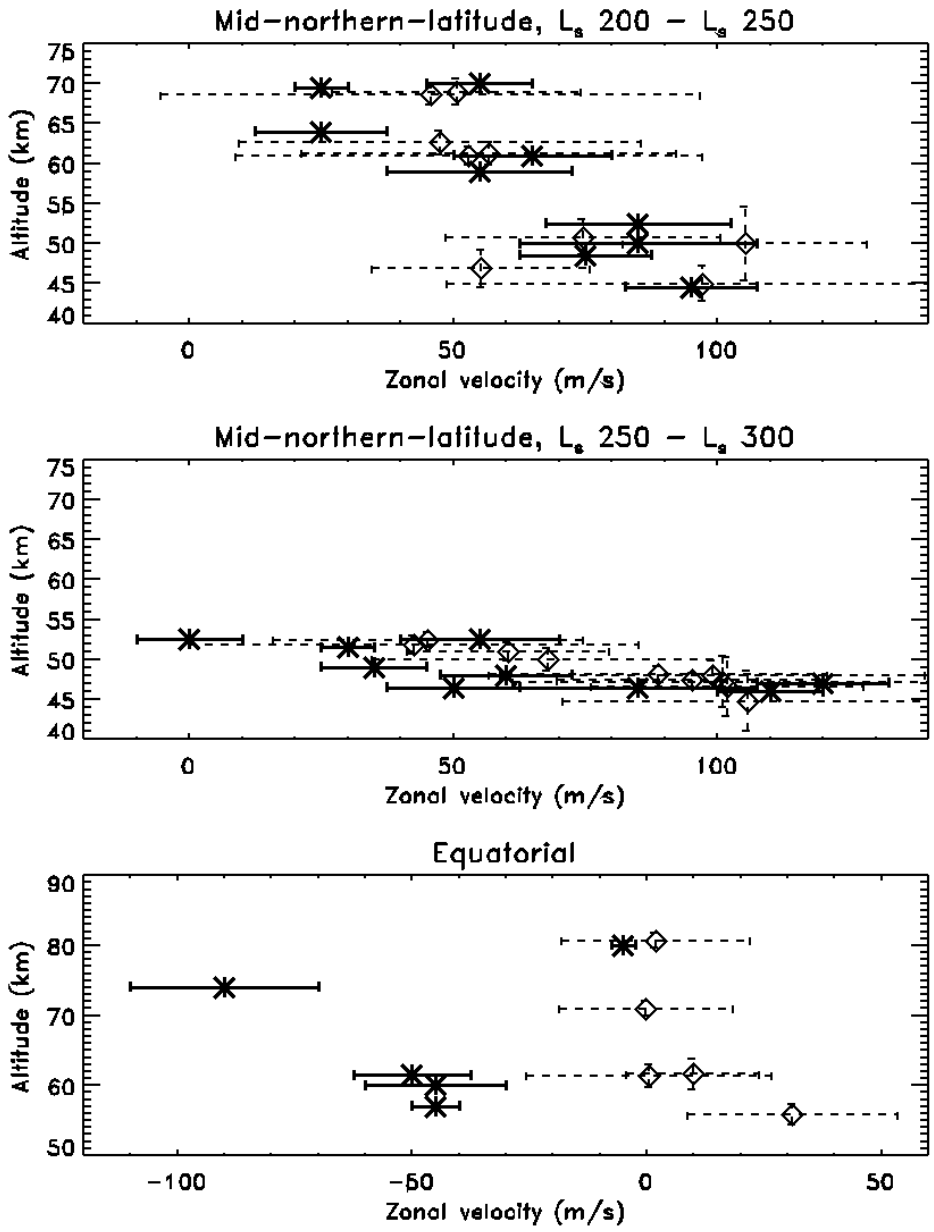

Figure 3.12: Comparison of THEMIS-VIS zonal velocity measurements with GCM values. Asterisks and solid lines show THEMIS-VIS measurements, diamonds and dashed lines show GCM values. The error bars correspond to our reported confidence intervals for the measurements, and to $\pm$ 2σ for the GCM values.(the standard error σ of the GCM values is defined in the text). We have suppressed the altitude error bars for the measurements, because they are quite small on this scale. For any given THEMIS-VIS image sequence, we show only one measurement and its corresponding GCM value. We chose the measurement with the narrowest confidence interval.

of the mesospheric clouds were detected in this L_s range. Once the typical incidence angles fell below 90°, no further mesospheric clouds were detected. All of the detected mid-latitude mesospheric clouds had 92° – 100° incidence, and none of the observations past L_s 300° had incidence angles in this range. Thus, the > 90° incidence angles may have been crucial to the detection of the mesospheric cloud features, but this is impossible to prove because of the correlation between seasons and incidence angle.

The distribution of mesospheric detections, as compared with all L_s 200° – 300° northern mid-latitude frontal storm survey image sequences having both blue and green filters, is shown in Figure 3.13. Each mid-latitude subclass is plotted with a different symbol. The distribution of mesospheric detections appears similar to the distribution of all sequences in the Acidalia region, with detections in 17% , 29 of 174 sequences. Only one out of 36 Amazonis sequences yielded a detection. This weakly suggest that either some intrinsic feature of the Acidalia region leads to more high altitude clouds, or that the slight difference in the latitude of the two survey regions has an impact on the presence or detectability of high altitude clouds.

Figure 3.14 shows several trends in the altitude of the mid-latitude mesospheric clouds. All of the clouds above 60 km altitude are found in L_s 200° – 240°, have incidence angles greater than 96°, and are in the "clumpy" morphology sub-class. Thus, it appears that later season, low-incidence, and linear or linear-periodic clouds are confined to lower mesospheric altitudes. However, since approximately three times as many incidence > 96° detections occur in L_s 200° – 240° as in L_s 240° – 300°, it is not clear whether the trends are related to the seasonal prevalence of certain types of clouds, or to their detectability as a function of incidence angle.

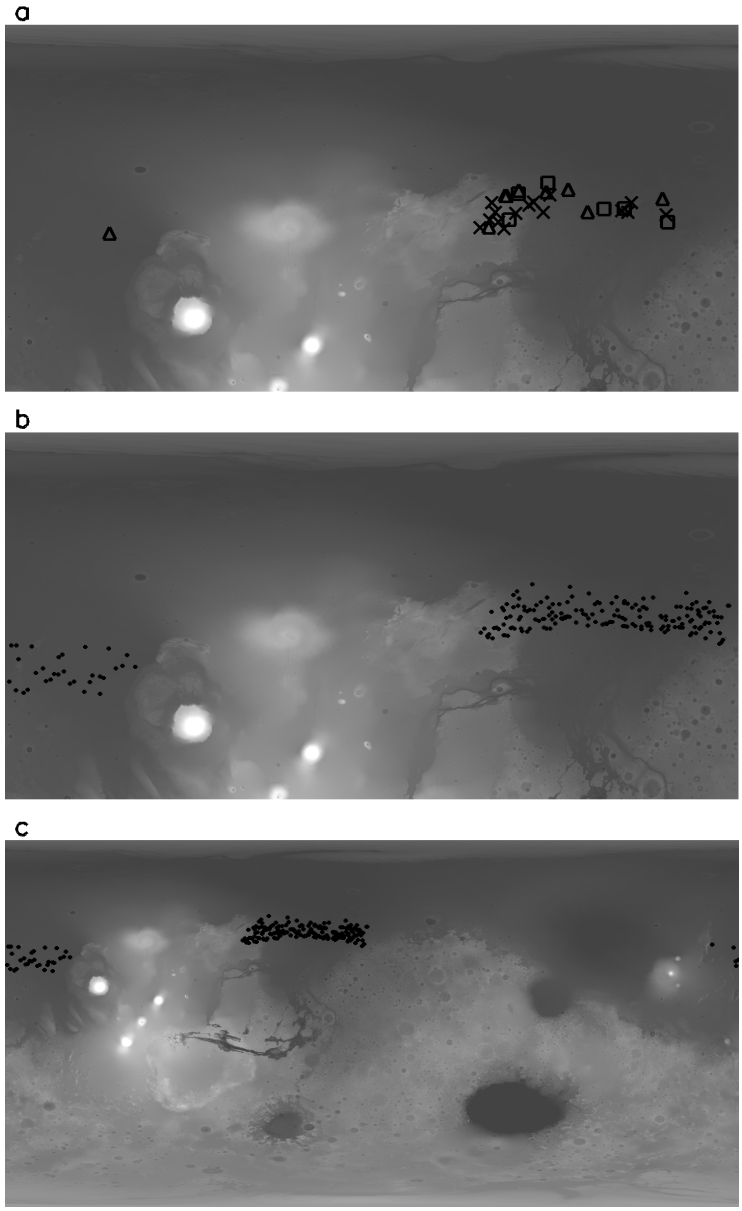

Figure 3.13: (a) Map of the locations where we have observed mid-latitude mesospheric clouds. Clouds in the "clumpy" subclass are marked with an "X", clouds in the "linear" subclass with a triangle, and clouds in the "linear-periodic" subclass with a square. In case where there are multiple mesospheric cloud ROIs in a single THEMIS-VIS image sequence, we have plotted only one from each subclass. (b) and (c) Location of all of the mid-latitude THEMIS-VIS images in the L_s 200° – 300° period that could have detected mesospheric clouds. (c) shows the entire planet. (b) shows the same region of the planet as in panel (a). The background for each map is grayscale-coded topography as measured by the Mars Orbiter Laser Altimeter.

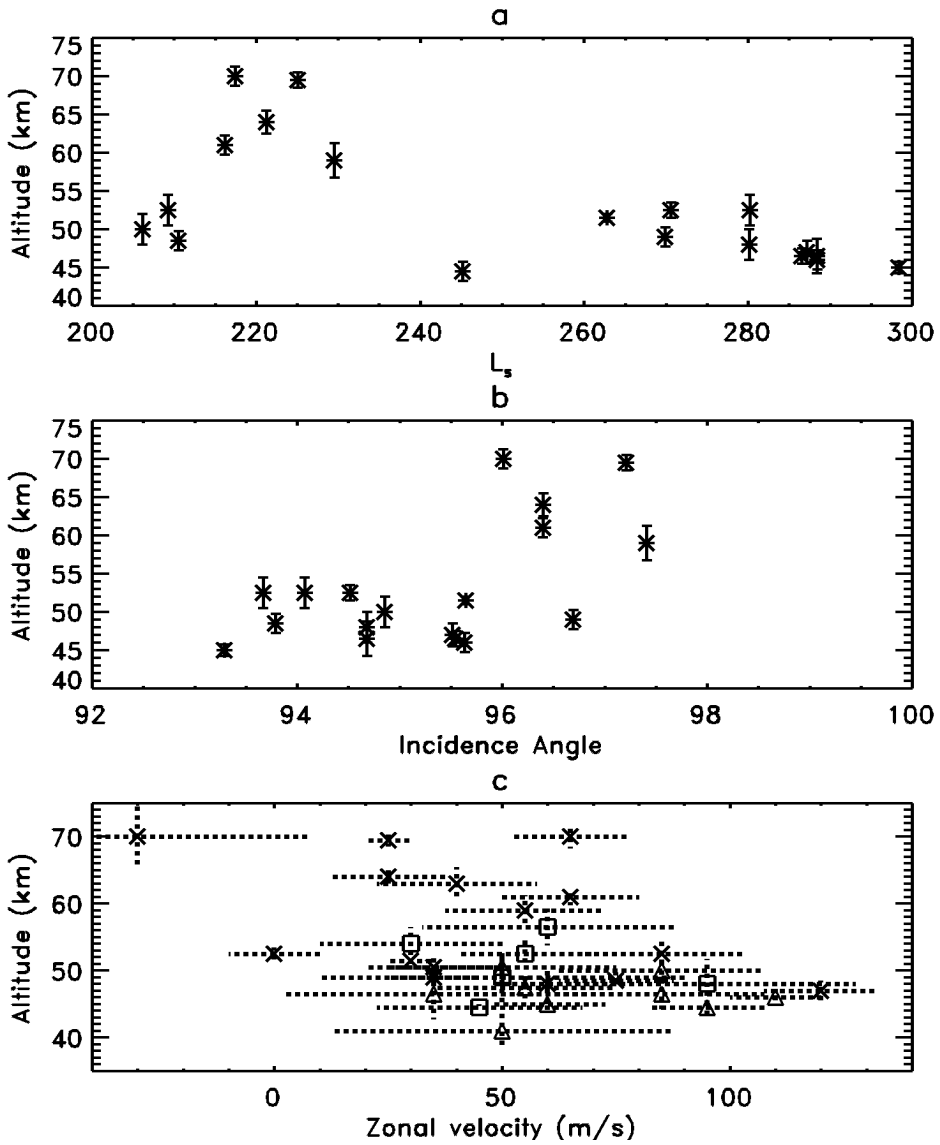

Figure 3.14: Altitude as a function of (a) season and (b) incidence angle for mid-latitude mesospheric clouds. (c) Altitude and velocity for the various subclasses of mid-latitude mesospheric clouds. Clouds in the "clumpy" subclass are marked with an "X", clouds in the "linear" subclass with a square, and clouds in the "linear-periodic" subclass with a triangle. For any given THEMIS-VIS sequence, we show only one measurement from each subclass. We chose the measurement with the narrowest confidence interval.

The incidence angles we report are for the planet's surface at the location of the mesospheric cloud correlation-ROIs, and of course high altitude clouds are still directly illuminated for surface incidence angles in the 90° – 100° range. Since we still see surface features at incidence angles greater than 95°, multiple scattering is obviously very important, and so we can't make simple assumptions about what altitude a feature must have to be detected. However, directly illuminated cloud features have a detectability advantage over lower altitude clouds, and given the observed correlations between incidence angle and cloud altitude, and the fact that we have not positively identified any low altitude clouds with incidence $> 92°$, our preferred hypothesis is that cloud aerosols are present at a range of altitudes during much of the northern fall-winter period, and the altitude of the aerosols that dominate the spatial contrast is selected by the geometry of the observation. Thus, when the incidence angle is low, the more diffuse high altitude features are overwhelmed by the lower altitude cloud, but at higher incidence angle the lower altitude aerosols contribute much less to the radiance reaching THEMIS-VIS, and so the higher altitude features become visible.

Comparison with Daily Global maps

In daily global maps (DGMs) of Mars assembled by Wang and Ingersoll [2002] from MOC-WA images, clouds are ubiquitous over northern mid-latitudes in the L_s 240° – 300° period. We have therefore compared our images of mesospheric clouds with the DGMs in an attempt to place them in context with the global cloud field. Although the "polar streak" clouds and polar lee waves described by Wang and Ingersoll [2002] are certainly prevalent in the general vicinity of the mesospheric cloud images, in almost all cases we cannot match any of the features

in the DGMs with features in the THEMIS-VIS images acquired on the same day. Since the wind speeds in this winter mid-latitude region of the mesosphere are high and the MOC-WA images are not concurrent with the THEMIS-VIS images, since THEMIS-VIS and MOC-WA observe at difference local times, and since the THEMIS-VIS field of view is very narrow compared to the MOC DGMs, it is not at all surprising that feature matching between the two instruments is difficult. It is also of course possible that some types of mesospheric clouds have no connection at all with the cloud field normally observed by MOC. However, we have identified two examples where an overlay of the THEMIS-VIS mesospheric image onto the MOC DGM seems to show a pattern similar to nearby MOC cloud features. These two examples are shown in Figures 3.15– 3.16. We don't expect an exact match because the two images are not concurrent, and so the feature match represents a subjective judgement. Accepting for the sake of argument that the match is real, we have a possible identification of the the linear mesospheric features with the "polar streak" clouds that are the dominant feature of the polar hood as described by Wang and Ingersoll [2002], implying that some of the familiar polar hood cloud extends up to 55 km in altitude.

Comparison with GCM

Both Tables 3.6 – 3.8 and Figure 3.12 compare our velocity measurements with WRF GCM results generated as previously described. Unlike the equatorial case, our mid-latitude velocities are entirely consistent with the GCM model predictions at all altitudes. Both the data and model output show a generally westerly (eastward) wind of the order tens, up to hundred, of m/s, and somewhat lower velocities above 55 km. Both the model and our measurements also show wide

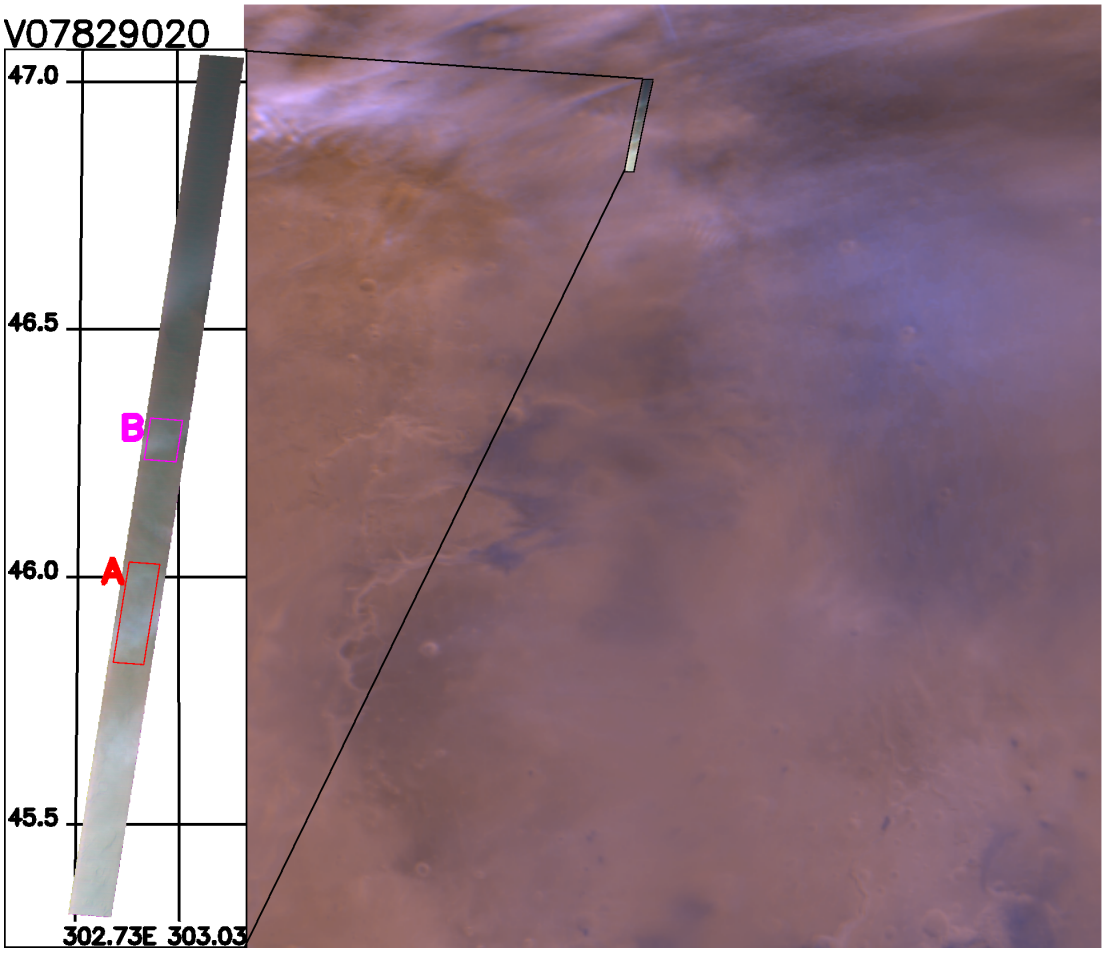

Figure 3.15: Comparison of a THEMIS-VIS high altitude cloud image with a portion of a MOC-WA DGM from the same day. The projection and labeling of the THEMIS-VIS image are the same as in Figure 3.7.

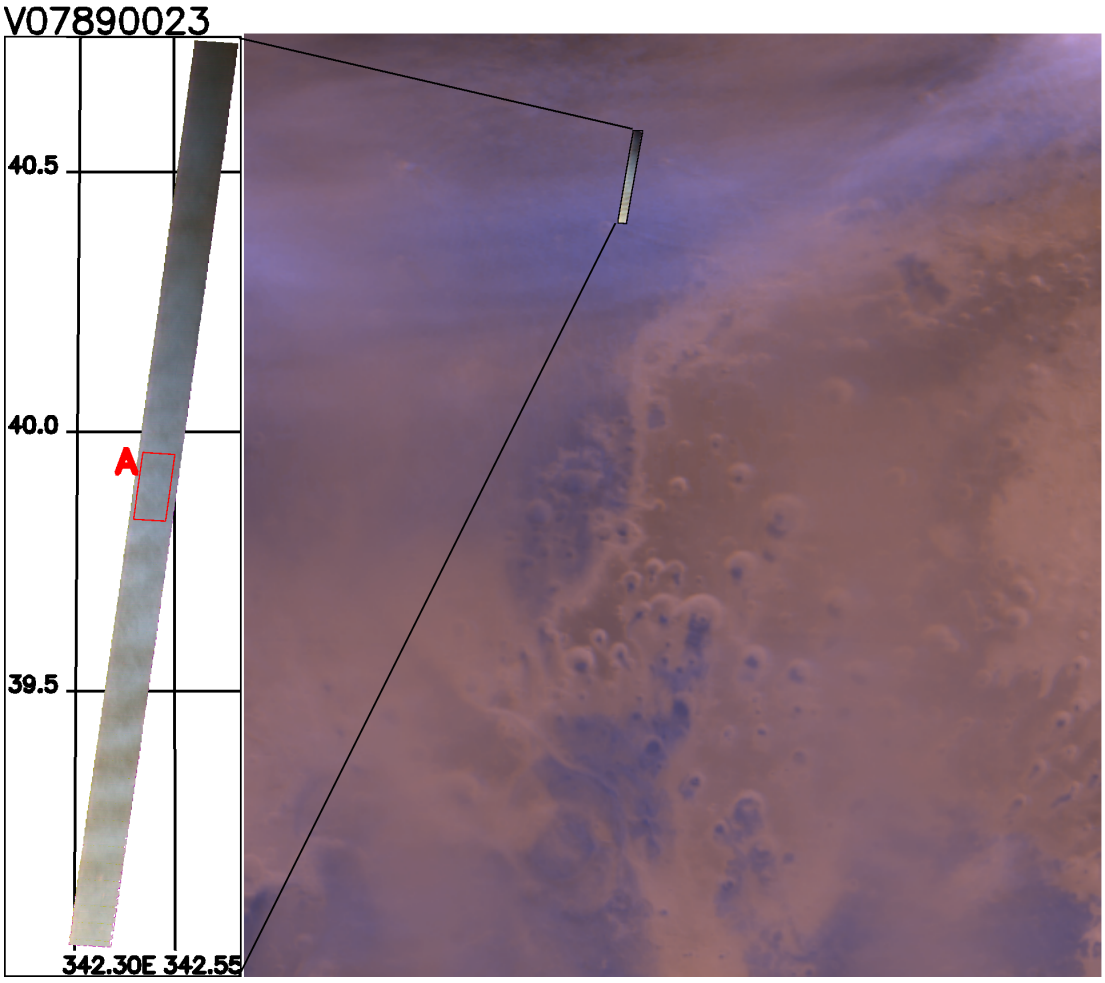

Figure 3.16: Comparison of a THEMIS-VIS high altitude cloud image with a portion of a MOC-WA DGM from the same day. The projection and labeling of the THEMIS-VIS image are the same as in Figure 3.7.

variability in the zonal wind at all levels, even to the point of occasional transient easterlies. This is consistent with the expected intense wave activity in the winter mid-latitudes. It also demonstrates that this eddy activity extends into the mesosphere where it might be important for the processes that form the observed clouds and cloud features.

3.4 Radiative transfer model

In order to compare our cloud measurements with previous work, and to provide insight into their physical significance, we apply a simple radiative transfer model to a subset of our data. We have chosen the standard, public domain, plain-parallel discrete ordinates model DISORT [Stamnes et al., 1988]. Prior uses of DISORT for martian clouds include Wolff et al. [1999] and Benson et al. [2003]. The choice of a plane parallel model immediately limits us to considering only the equatorial clouds, which are observed at incidence angles significantly less than 90 degrees. Investigation of the properties of the mid-latitude twilight clouds will have to wait for a more sophisticated 3-dimensional modeling approach, which is beyond the scope of this work.

We apply our DISORT model to portions of three equatorial mesospheric cloud images, We begin by drawing a pair of regions of interest on the image; one region with minimal blue filter brightness and thus apparently minimal cloud opacity; and one region with maximal blue filter brightness and thus maximal cloud. These regions are shown in Fig. 3.7. For V04573003 and V11100003, we have selected two such pairs, in order to check the consistency of our model results in different portions of the image. V10526009 has only a small area of mesospheric cloud, and so we have defined only one region of interest pair. For each filter, we take

the mean radiance in the selected regions, labeling them the "cloud radiance" and the "cloud-free radiance." The radiance calibration standard errors, reported in chapter 4, are much larger than the standard deviations of the radiance in our selected regions, and so we adopt the chapter 4 standard errors.

The models include a lambertian surface as the lower boundary, rayleigh scattering by gaseous atmospheric constituents, a lower-atmosphere aerosol component with properties consistent with previous work, and a mesospheric aerosol component. Our DISORT models use 64 streams and 8 vertical layers, each 10 km thick, with the bottom of the lowest layer being at the surface and the top of the top layer at 80km. For each layer we specify that layer's optical depth, single-scatter albedo (SSA), and phase function. The optical depth of a layer is a sum of the optical depths of whichever of the three components are present in the layer. The SSA and phase functions are weighted averages of the component values, where the weights are the component optical depths in that layer. The rayleigh scattering is determined by a density profile that we derive via the hydrostatic equation from a typical low latitude MGS-TES temperature profile [*e.g.*, Conrath et al., 2000]. We adjust the surface pressure to account for the altitude of the local surface in each model.

3.4.1 Lower atmosphere and surface

MGS-TES [*e.g.*, Smith et al., 2001] measurements show very low dust optical depths at the time and place of all of the observations we model, and so the lower aerosol component is defined to be consistent with Clancy et al. [2003] "type 1" water-ice aerosols. Thus, we use the Clancy et al. [2003] measured phase function; adopt a particle size distribution with 1.5 μm effective radius (r_{eff}) and an

effective variance (v_{eff}) of 0.1 for purposes of scaling the opacity as a function of wavelength; and assume zero water ice aerosol below 20 km altitude. Above 20 km, we assume that the water ice aerosol opacity in any given layer scales with the gas density, up to an arbitrary upper boundary of 60 km. We have found that the choice of this upper boundary has no effect on our results.

The normal albedo of the surface and the optical depth of the lower aerosol layer are constrained using our cloud-free radiance measurement and "nearby" MGS-TES aerosol optical depth measurements. We presume that, since the mesospheric aerosols are invisible to THEMIS-IR, they are also invisible to MGS-TES. For the case of V10526009, we defined "nearby" as within 3 sols of time, 5° of longitude and 0.5° of latitude, and adopted the mean of TES aerosol retrievals in this region. For V11100003, we instead used 1 sol and 1.5° of latitude as the limits. THEMIS-IR aerosol retrievals [Smith et al., 2003] are available concurrent with V11100003, and these measurements are consistent, within the uncertainties reported by Smith et al. [2003], with our adopted MGS-TES values. In the vicinity of V04573003, no TES or THEMIS-IR aerosol retrievals are available, due to low surface temperatures.

Since the TES retrievals consider only absorption, they need to be adjusted to provide an estimate of the full extinction optical depth. According to Smith [2004], the appropriate adjustment is to multiply the TES optical depth by 1.5. We also of course must adjust the TES optical depth to compensate for the difference in extinction cross section between the wavelength where it is defined (12.1 μm), and the wavelengths of the THEMIS-VIS filters. We calculate the cross sections as a function of wavelength using standard Mie theory *[e.g. Bohren and Huffman, 1983]* as implemented in the publicly available code, DMiLay (`ftp://climate.gsfc.`

nasa.gov/pub/wiscombe/Single_Scatt/Coated_Sphere). The water ice indices of refraction (real and imaginary) are from Warren [1984].

Where the TES retrievals are available, we use the TES derived optical depth together with the above described assumptions about the aerosol properties, and iteratively adjust the surface albedo until the DISORT model matches the cloud-free radiance. We perform this procedure independently for each THEMIS-VIS filter, and, for simplicity and efficiency, we model only the single discrete wavelength at the center of the bandpass for each filter. Since the assumption that the surface is a lambertian scatterer is probably a poor one, we do not expect to obtain a normal albedo that is consistent with previous work or even representative of the real normal albedo of the surface. Rather, we obtain an effective albedo that happens to be appropriate for the high incidence angle geometry of the THEMIS-VIS measurements. Estimates of the true photometric properties of the surface at high incidence angle are not available, and so solving for an effective albedo represents a best guess. More importantly, the function of the albedo in our model is merely to allow the surface / lower atmosphere components of our model to match the observed cloud-free radiance. Since we always achieve such a match, the primary effect of trading off albedo for lower-atmosphere optical depth is to alter the angular distribution of radiation reaching the mesospheric region from below. This angular distribution has only a secondary effect on our model. We have experimented with decreasing the lower-atmosphere opacity by a factor of two, which leads to unrealistically high surface albedos, but has negligible impact on our derived mesospheric optical depths and the perturbation to our derived particles sizes is small compared with other sources of uncertainty.

Since V04573003 lacks TES or THEMIS-IR measured optical depths, we follow

a slightly more complicated procedure for that case. We start with the blue filter normal albedo derived for V10526009, then iterate the blue filter optical depth to match the blue cloud-free radiance. Since the blue filter albedo of the martian surface is very low, its influence on the observed radiance is small when the optical depth is high, as it is for V04573003. Having derived the blue filter optical depth, we simply scale it to the other filters using the Mie-derived cross sections, and then iterate for the albedo in these filters as previously described.

For each pair of radiance measurements, we crucially assume that both the surface albedo and the lower-atmosphere aerosol component are the same in the (mesospheric) cloud-free region and the mesospheric cloud region. The mesospheric clouds in V12922001 and V13072001 are in regions of prominent surface features, which is why we have not attempted to apply our model to them. For the image sequences that we do model, the cloud and cloud-free regions are drawn to minimize the possibility of such surface variations, and it is clear from Fig. 3.7 that the mesospheric aerosols are the dominant source of contrast in their vicinity. Since Fig. 3.7 is projected at the measured aerosol altitude, significant surface contrasts lead to obvious color mis-alignments, as would low-altitude aerosol variations if they were present.

The other crucial assumption that we make about the relationship between the cloud-free radiance and the cloud radiance is that the cloud-free region actually contains zero mesospheric aerosol, rather than simply less mesospheric aerosol than the cloud features themselves. We have no way to verify this assumption, but it is necessary, because without it we have no way to constrain the effective surface albedo. If there is a mesospheric aerosol contribution to the radiance that we have labeled cloud-free, it would mean that we have overstated the lower

atmosphere / surface contribution to the observed radiance in the cloudy areas, and understated the mesospheric contribution to the radiance in these areas. Since we are working in a regime with a positive relationship between optical depth and observed radiance, any mesospheric aerosol contamination of the "cloud-free" areas leads us to underestimate the mesospheric aerosol optical depth.

3.4.2 Mesospheric aerosols

Having used the above assumptions to establish the contributions of the surface and lower atmosphere, we explore a range of possibilities for properties of the mesospheric clouds. We must first specify the upper and lower altitude boundaries of the mesospheric (we choose 70 – 80 km), but we have found these boundaries have no effect on the modeled radiances, and so they are extraneous to the modeling problem. The factors that do influence the radiance we observe, as modeled by DISORT, are the vertical extinction optical depth, the single scatter albedo (SSA), and the single-scatter phase function of the aerosols, all of which themselves are functions of wavelength. Since this modeling problem is highly under-determined, we first simplify by adopting a Henyey-Greenstein phase function [Henyey and Greenstein, 1941] for the mesospheric aerosols, which can be described by a single parameter, g (the asymmetry parameter), at each wavelength. Next, we impose a physical description of the aerosol particles, which gives us, via Mie theory, g_λ, SSA_λ, and the mean particle extinction cross section $\langle\sigma\rangle_\lambda$, at the wavelengths λ of all of the THEMIS-VIS filters. Now, the problem is overdetermined for any given physical description — specifying the optical depth in any one filter yields the optical depth in the rest, via their relative values of $\langle\sigma\rangle_\lambda$, and so a single optical depth yields radiance in all three or four filters. We arbitrarily choose the blue

filter as the wavelength at which to specify optical depth. Treating the observed cloud radiances as random variables with standard errors equal to the THEMIS-VIS calibration standard errors, we have a χ^2, and so we find the blue filter optical depth by iteratively adjusting it until the χ^2 is minimized. The probability of observing a χ^2 as large as the minimized χ^2 can be interpreted as a likelihood for the best fit with a particular physical description of the aerosols. Thus, we can reject any physical description that yields a best-fit likelihood lower than some arbitrary cutoff.

The aerosols that we have chosen to consider for the mesospheric clouds are:

1. CO_2 ice [indices of refraction from Hansen, 1997]
2. Water ice [indices of refraction from Warren, 1984]
3. Dust [indices of refraction from Wolff and 11 co-authors, 2006]
4. Dust core with CO_2 ice shell: core radii 20%, 40%, 60%, and 80% of the total particle radius
5. Dust core with water ice shell: core radii 20%, 40%, 60%, and 80% of the total particle radius

For each aerosol type, we consider particle sizes ranging from $r_{\text{eff}} = 0.05$ μm to $r_{\text{eff}} = 5$ μm. We have chosen these limits because in most cases they appear to bound the region where significant model likelihood is observed. We test particle size with sampling intervals of 0.01 μm between 0.05 and 0.2, 0.1 between 0.2 and 2.0, and 0.5 from 2.0 to 5.0. For all cases, we use a simple shape for the mesospheric particle size distribution: a gamma distribution with v_{eff} of 0.2, consistent with Chassefière et al. [1992]. An experiment with $v_{\text{eff}} = 0.1$ showed negligible changes in our results.

The radiances, geometric parameters, and assumed surface and lower-atmosphere properties used for our models are shown in Table 3.9. All models of a given ROI pair use the same assumptions for the surface and lower atmosphere. We construct separate models with optimized optical depth for each particle composition and particle size. In the results tables (Tables 3.10–3.11) we show the particle size, optical depth, and, where applicable, the core radius for the model with the highest likelihood. We also show confidence intervals for these parameters that include all models with likelihood greater than our chosen threshold of 1%. Particle compositions that yield no model likelihoods greater than this threshold are indicated with a '–' in the table.

For V04573003 and V11100003, we have included in the results tables only the ROI pair with the highest model likelihood from each image sequence. In both of the sequences with two ROI pairs, both ROI pairs yielded similar likelihood maxima with similar particle properties. The confidence intervals are listed as "any" for V10526009 because for this image sequence all the models yielded likelihoods greater than the cutoff. The lack of precision with this image is a result of its short exposure time, which leads to high uncertainty in the radiance calibration, and of the low contrast of its aerosol features.

Since 4.0μm particle size was the upper bound of the searched parameter space, a value of 4.0μm for the upper limit means that no upper bound of the particle size was found. When the particle size confidence interval is listed as "any," or has an upper limit of 4.0μm, the τ confidence interval simply reflects the range of values observed in the searched parameter space, and so is not a true constraint on the optical depth. However, in our model the largest particle sizes yield the largest optical depths, so when there is a valid lower bound on the particle size, the

Table 3.9: Radiative Transfer ROIs: mean properties of each cloud / cloud-free pair

Property	ROI pair				
	V04573003		V10526009	V11100003	
	J, K	X, Y	J, K	J, K	X, Y
Cloud-free ROI viewing geometry					
E. lon.	261.172	261.207	261.929	290.91	290.842
N. lat.	−16.264	−16.623	0.536	−13.98	−14.283
° Inc.	80.02	80.24	68.96	77.89	77.96
Cloud ROI viewing geometry					
E. lon.	261.209	261.235	261.966	290.95	290.95
N. lat.	−16.266	−16.619	0.522	−14.05	−14.306
° Inc.	80.06	80.26	69.00	77.95	78.06
Cloud-free radiance $((I/F) \times 10^{-2})$					
at 0.425 μm	2.86	2.67	3.73	2.00	2.16
at 0.540 μm	5.11	4.61	7.30	3.34	3.39
at 0.654 μm	8.27	7.49	14.8	–	–
at 0.749 μm	9.90	8.97	18.4	6.03	5.96
Cloud radiance $((I/F) \times 10^{-2})$					
at 0.425 μm	3.40	3.19	4.93	2.95	3.22
at 0.540 μm	5.59	5.15	7.75	4.29	4.47
at 0.654 μm	8.72	8.15	15.1	–	–
at 0.749 μm	10.4	9.71	18.6	7.18	7.45
Lower-atmosphere aerosol component optical depth, τ_{low}					
at 0.425 μm	0.54	0.49	0.44	0.23	0.23
at 12.1 μm	0.28	0.25	0.23	0.12	0.12
Effective surface albedo, A_{eff}					
at 0.425 μm	0.023	0.023	0.020	0.023	0.034
at 0.540 μm	0.271	0.237	0.16	0.12	0.12
at 0.654 μm	0.579	0.522	0.43	–	–
at 0.749 μm	0.727	0.661	0.55	0.29	0.284

Table 3.10: Radiative Transfer model results: homogeneous mesospheric cloud aerosols

	V04573003b	V10526009	V11100003	V11100003 749 nm excluded
ROI pair	J, K	J, K	X, Y	X, Y
		CO_2 ice		
r_{eff} (μm)	0.10 [0.08 – 0.11]	0.05 [any]	–	1.5 [0.8 – 1.8]
τ at 0.425 μm	0.22 [0.14 – 0.37]	0.074 [0.074 – 0.31]	–	0.50 [0.38 – 0.57]
τ at 9.3 μm	0.00	0.00	–	0.05
τ at 12.1 μm	0.00001	0.00001	–	0.02
n (cm^{-3}) [1]	2080	19400	–	6.42
Mass fraction [2]	0.07%	0.08%	–	0.07%
Precipitable μm [3]	0.04	0.05	–	0.44
		H_2O ice		
r_{eff} (μm)	0.10 [0.08 – 0.12]	0.05 [any]	–	–
τ at 0.425 μm	0.22 [0.13 – 0.38]	0.073 [0.073 – 0.44]	–	–
τ at 9.3 μm	0.004	0.005	–	–
τ at 12.1 μm	0.03	0.03	–	–
n (cm^{-3}) [1]	3710	34700	–	–
Relative mass [4]	6.5	7.6	–	–
Precipitable μm [3]	0.07	0.09	–	–
		Dust		
r_{eff} (μm)	0.08 [0.07 – 0.09]	0.10 [0.05 – 0.4]	0.5 [0.5 – 0.5]	0.4 [0.4 – 0.5]
τ at 0.425 μm	0.30 [0.23 – 0.40]	0.20 [0.13 – 0.94]	1.6 [1.6 – 1.6]	1.0 [1.0 – 1.6]
τ at 9.3 μm	0.04	0.02	0.29	0.13
τ at 12.1 μm	0.004	0.002	0.03	0.013
n (cm^{-3}) [1]	5350	1380	153	141

[1]Number density, assuming that the physical depth of the cloud is 10 km.

[2]Mass of CO_2 condensate as a fraction of the mass of the atmosphere between 70 and 80 km (for V04573003 and V10526009), or between 50 and 60 km (for V11100003).

[3]Depth of condensate if spread uniformly in solid phase on a horizontal surface.

[4]Ratio of the mass of H_2O condensate to the mass of H_2O vapor in a 10 km, saturated column, at 160 K.

Table 3.11: Radiative Transfer model results: heterogeneous mesospheric cloud aerosols

	V04573003b	V10526009	V11100003	V11100003 749 nm excluded
ROI pair	J, K	J, K	X, Y	X, Y
		Dust core, CO_2 ice shell		
r_{eff} (μm)	0.09 [0.08 – 0.11]	0.05 [any]	1.4 [0.7 – 4.0]	0.7 [0.5 – 4.0]
Core radius fraction	0.8 [0.2 – 0.8]	0.2 [any]	0.6 [0.4 – 0.8]	0.8 [0.2 – 0.8]
τ at 0.425 μm	0.25 [0.14 – 0.41]	0.075 [0.075 – 23]	1.3 [0.75 – 2.8]	1.1 [0.38 – 23]
τ at 9.3 μm	0.016	0.0003	0.30	0.17
τ at 12.1 μm	0.015	0.00004	0.08	0.02
τ_{dust} at 9.3 μm [1]	0.014	0.0003	0.15	0.15
n (cm^{-3}) [2]	3040	19600	18.4	55.5
Mass fraction [3]	0.04%	0.08%	0.1%	0.03%
Precipitable μm [4]	0.02	0.08	0.8	0.2
		Dust core, H_2O ice shell		
r_{eff} (μm)	0.10 [0.08 – 0.12]	0.05 [any]	0.8 [0.8 – 3.5]	0.7 [0.6 – 4.5]
Core radius fraction	0.6 [0.2 – 0.8]	0.2 [any]	0.8 [0.4 – 0.8]	0.8 [0.4 – 0.8]
τ at 0.425 μm	0.25 [0.13 – 0.41]	0.073 [0.073 –1.8]	1.4 [1.4 – 2.7]	0.97 [0.79 – 61]
τ at 9.3 μm	0.013	0.005	0.29	0.17
τ at 12.1 μm	0.019	0.03	0.15	0.09
τ_{dust} at 9.3 μm [1]	0.009	0.0003	0.25	0.14
n (cm^{-3}) [2]	3080	34100	59.3	53.1
Relative mass [5]	4.2	7.4	26	15
Precipitable μm [4]	0.05	0.09	0.3	0.2

[1]The vertical optical depth that the dust cores would have if not coated by condensate.

[2]Number density, assuming that the physical depth of the cloud is 10 km.

[3]Mass of CO_2 condensate coatings as a fraction of the mass of the atmosphere between 70 and 80 km (for V04573003 and V10526009), or between 50 and 60 km (for V11100003).

[4]Depth of condensate if spread uniformly in solid phase on a horizontal surface.

[5]Ratio of the mass of the H_2O condensate coatings to the mass of H_2O vapor in a 10 km, saturated column, at 160 K.

lower bound on the optical depth can also be considered valid. Once the optical depth is larger than ∼5, changing it has very little effect on the modeled radiance, and so optical depth limit values greater than 5 are numerical artifacts and can be interpreted simply as "very large."

For the best fit particle size, optical depth, and core radius, we present various other quantities intended both as reality checks and as indicators of the physical significance of the results. In order to calculate number density we have assumed a physical depth for the clouds of 10 km. For the mass fraction of CO_2, we must chose a portion of the atmospheric column to which to compare the column mass of condensate; we have chosen a 10 km vertical column immediately below the measured cloud altitude, determining its mass from the same MGS-TES-derived density profile used to calculate Rayleigh scattering. For the relative mass of H_2O, we compare the mass of H_2O in the aerosols with the maximum mass of H_2O vapor (*i.e.*, mass at saturation) that could be contained in a 10 km vertical column at a reasonable upper bound [see Clancy and Sandor, 1998] temperature for the mesosphere of 160 K. With $\tau_{\rm dust}$ we suppose that the refractory cores of the heterogeneous aerosols must have been present before the condensate shells formed, and calculate the optical depth that they would have had before the condensation occurred.

3.4.3 Constraints imposed by THEMIS-IR and MGS-TES

Both V04573003 and V11100003 were accompanied by simultaneously acquired THEMIS-IR images. Neither IR image shows any evidence for cloud features, which provides a powerful constraint on the aerosol composition. Also, if we suppose that the dust cores of the heterogeneous aerosols must be present over a

somewhat wider region than the locus of condensation itself, the very low dust optical depths measured by MGS-TES also provide a constraint. Since THEMIS-IR and MGS-TES routinely measure dust and water-ice optical depth using absorption bands centered at 9.3 and 12.1 μm, respectively, [Smith et al., 2001; Smith, 2004] we calculate optical depths at these wavelengths to see if either instrument would have detected the aerosol in question.

Our models for V04573003 and V10526009 (considering the best fit) yielded particle sizes and optical depths small enough to be invisible to the infrared instruments. However, both THEMIS-IR and MGS-TES arguments rule out the dust and the heterogeneous aerosols for V11100003. THEMIS-IR measures dust optical depth with an uncertainty of about 0.04 [Smith, 2004], and all of these cases in V11100003 yield 9.3 μm optical depths greater than 0.17. Smith et al. [2001] give 0.05 as the uncertainty for MGS-TES optical depths, and the mean MGS-TES dust optical depth in the vicinity of V11100003 was 0.03, so the the τ =0.14 of dust optical depth required as condensation nucleii for the aerosols in our model are unlikely to be present, even if we supposed that all of the dust aerosol was (improbably) in the mesosphere.

3.4.4 Models with the 749 nm filter excluded

Ruling out the heterogeneous aerosols obviously poses a problem for the first V11100003 column in the results table in that it eliminates all of the models that exceed our likelihood threshold. This motivates us to consider models that ignore the 749 nm filter radiances, which are shown in the final column tables Tables 3.10–3.11. Since the martian surface albedo is highest in the 749 nm filter, its influence on the models is also greatest in the 749 nm filter. Thus the assumption of constant

surface albedo between the cloud and cloud-free radiance ROIs is most critical at this wavelength. Although we strove to minimize the chance of albedo contrasts, V11100003 has more prominent surface features than the other images, and so albedo variations are a plausible explanation for lack of tenable results when the 749 filter is included. For the sake of consistency, we have experimented with neglecting the 749 nm filter in the V04573003 cases and found no change in the best fits or the confidence intervals. In fact, neglecting *both* the 654 nm filter and the 749 nm filters also yields no change in the V04573003 best fits, and causes only a slight widening of the confidence intervals.

3.4.5 Interpretation

Given the simplifications and assumptions we have made, our model is hardly unique, but we believe that it includes as much detail as can be supported by our observations. Within the framework of our assumptions, two main conclusions can be drawn:

1. The V04573003 cloud has 0.1 ± 0.02 μm aerosols, in good agreement with Clancy and Sandor [1998], and optical depth much larger than that of previously reported mesospheric hazes and twice as large as that proposed by Clancy and Sandor [1998] for their Pathfinder clouds. The complex cloud morphology tends to imply that a condensate is involved. Furthermore, since the mass of condensate is at least four time greater than the maximum amount of water vapor that can be contained in a mesospheric column, most of the condensate is likely to be CO_2.

2. The V11100003 cloud has large $\sim$1.5 μm particles and a very substantial

optical depth of ∼0.5. Both these values are comparable to lower atmosphere water ice clouds. Our radiative transfer models weakly prefer CO_2 over water ice as the composition of these clouds. Even if our models did admit water condensate, the mass of water required would be much higher than the carrying capacity of the mesosphere, and so we have a strong argument for most of the condensate being CO_2. Any condensation nuclei present in these clouds must be a small fraction, 20% or less, of the particle radii, otherwise it would have been detected by one of the infrared instruments.

The diversity in particle size is consistent with the diversity observed by Chassefière et al. [1992]. Despite this diversity, both clouds, perhaps coincidently, contain condensate equal to 0.07% of the mass of the atmosphere in the 10 km layer below the altitude of the observed features. This means that the V11100003 cloud has about ten times as much mass per unit area because it is 2 scale heights below V04573003. Both clouds have a 0.1 μm upper limit on the size of the condensation nuclei, which means that, according to Colaprete and Toon [2003] (their Figure 3), mesospheric temperatures must have transiently reached as low as 90 K near 80 km altitude in order to initiate condensation. This is 5 K below the coldest temperatures observed by Pathfinder, and substantially colder than any other mesospheric temperature measurements [Clancy and Sandor, 1998].

3.5 Conclusions

THEMIS-VIS was not designed for stereo capability, nor intended to study mesospheric clouds, but nevertheless it can be used to measure mesospheric cloud altitude and zonal velocity. Additionally, the THEMIS-VIS multispectral capability and radiance calibration allows it to additionally place constraints on the aerosol

properties of the observed cloud. The altitude and velocity measurements have precisions as good as ± 1 km and ± 10 m/s, respectively. Systematic errors in the THEMIS-VIS camera model add an additional ± 1 km and ± 10 m/s to the uncertainties, and any meridional velocity will produce an error of about 1 km in altitude for every 10 m/s.

Using this capability, we have presented the first horizontally resolved images of verifiably mesospheric clouds, and find that they fall into two distinct classes in the THEMIS-VIS data set: equatorial mesospheric clouds, and northern mid-latitude winter mesospheric clouds. The existence of the equatorial mesospheric clouds was suggested by Clancy and Sandor [1998]. We have confirmed the mesospheric altitude of such clouds, and identified one example that has an altitude, particle size, and optical depth all consistent with the Clancy and Sandor [1998] explanation for the Pathfinder blue wave clouds. The optical depths of these discrete clouds are an order of magnitude higher than that of the limb hazes observed by Viking and MOC, and so the prevalence of these clouds becomes an important question for both the chemistry and radiative balance of the mesosphere. The low detection rate in the THEMIS-VIS data set perhaps implies that they are rare and thus unimportant, but other factors besides scarcity of clouds, such as the time of day of the ODY orbit and the sensitivity of THEMIS-VIS will need to be considered as part of a careful statistical analysis in order for their true impact to be understood.

We are aware of no previous predictions of discrete mesospheric cloud features in winter mid-latitudes. Their existence places an important new constraint on the cloud formation mechanism or mechanisms that operate in the vicinity of the polar hood, in that they must be operating at altitudes up to 70 km. The observed horizontal, vertical, and temporal distribution of the clouds is most likely

controlled by the peculiar pattern of the observations that led to their serendipitous discovery, and thus we have almost no information about the real distribution of this type of cloud, other than that they appear to be visible only in twilight, and, with a detection rate of 17% in Acidalia, fairly common. We do, however, see some evidence that identifies the lineated mesospheric clouds that have 40 - 55km altitudes with the polar streak clouds that are ubiquitous in the polar hood.

There is no reason to believe that the mid-latitude mesopheric clouds and the equatorial mesospheric clouds have anything in common besides the manner in which they were detected. For the equatorial clouds, the Clancy and Sandor [1998] proposal of gravity wave induced supersaturation leading to CO_2 cloud formation is quite consistent with our observations. The mass of aerosol is too large to be water. Furthermore, the wave-like morphology of the clouds and the observed size of the aerosols points to a coherent story for CO_2 aerosol formation that requires gravity wave amplitudes only modestly greater than those observed during the Pathfinder descent.

We have little information on the composition of the mid-latitude mesospheric clouds. The cloud morphologies that we show here may or may not provide insight into their mechanism of formation, and we must of course consider that the clouds at 70 km in altitude may be different in nature than those at 45 km. Given that the mid-latitude winter mesospheric temperatures are probably warmer than those of the tropical mesosphere [Smith et al., 2001; Joshi et al., 1995], these clouds are less likely to be CO_2. Since our modeling of THEMIS-VIS images yielded meaningful constraints on the nature of the equatorial clouds, it seems likely that a radiative transfer model suitable for twilight will ultimately yield similar constraints on the mid-latitude clouds.

The poor match between our equatorial mesospheric velocities and our GCM results provides an important challenge for future modeling work. Ideally, these models would provide a description of gravity wave drag which produces both the correct wind speeds and gravity wave amplitudes sufficient to produce the necessary supersaturation at the altitudes that we observe clouds. We might also hope that such a model would explain the apparent longitudinal asymmetry in the cloud observations. A still more ambitious theoretical or modeling question is the relationship between gravity wave breaking and mesospheric clouds morphology. What would wave-breaking turbulence in mesopheric clouds look like, and is it consistent with our observed morphologies?

A crucial future direction for observational study of these mesospheric clouds is a detailed examination of MOC-WA limb observations with the goal of identifying the relationship between the limb hazes and the mesospheric clouds. Similar comparisons will be possible with MARCI [Malin et al., 2001] on Mars Reconaissance Orbiter. MARCI, a wide-angle (180°) multi-spectral imager with 2 UV channels plus visible coverage similar to THEMIS-VIS, may also be able to provide constraints on the limb-haze composition and particle size.

As long as THEMIS-VIS continues routine imaging of martian surface features using the green and blue filters, it is likely to continue to discover equatorial mesospheric clouds. The relative allocation of THEMIS-VIS imaging between multispectral and monochromatic sequences will obviously greatly affect the rate of detection. Future THEMIS-VIS observations of mid-latitude mesospheric clouds depends on the details of the ODY orbit during the next northern fall-winter period. If the orbit puts ODY over twilight surfaces in northern fall-winter mid-latitudes, we recommend regular twilight imaging with uniform sampling in lon-

gitude in order to determine the true spatial distribution of the clouds. We also recommend using four filters instead of three to provide more robust constraints on radiative transfer models, and 4x4 binning to maximize coverage and thus the odds of a detection. If spacecraft engineering constraints allow for off-nadir pointing of THEMIS-VIS, then a mixture of twilight and pre-sunset images should be acquired at any given latitude in order to clarify the effect of incidence angle on the clouds or on their detectability.

The THEMIS-VIS altitude-velocity measurement capability is not in principle limited to mesospheric clouds. At lower altitudes, however, identification of cloud features becomes more problematic, and the confounding effect of nearby surface features on our cross-correlation method becomes more significant. Thus, future work in this direction will involve the development of more sophisticate techniques for feature localization.

CHAPTER 4

CALIBRATION AND IN-FLIGHT PERFORMANCE OF THE MARS ODYSSEY THEMIS VISIBLE IMAGING SUBSYSTEM (VIS) INSTRUMENT

4.1 Introduction

The Mars Odyssey spacecraft's Thermal Emission Imaging System (THEMIS) consists of two independent multispectral imagers sharing a single telescope — a nine-band mid-infrared microbolometer arrary (THEMIS-IR), and a five-band visible/near-infrared interline-transfer CCD (THEMIS-VIS). THEMIS has been acquiring visible and infrared images from Odyssey's 400 km circular polar orbit since February 2002, and has already provided important new insights about the geology and evolution of the martian surface. [*e.g.*, Christensen et al., 2003; Titus et al., 2003; Pelkey et al., 2003, 2004; Milam et al., 2003; Bell, 2003] Details of the THEMIS design, including THEMIS-IR, and the optics shared by both subsystems, can be found in Christensen et al. [2004]. The purpose of this chapter is to describe and evaluate the calibration of THEMIS-VIS, and to describe aspects of the instrument design and on-orbit performance which are directly related to the calibration.

THEMIS-VIS is one of several visible light instruments currently in operation in Mars orbit. Of these, the Mars Orbiter Camera (MOC) on the Mars Global Surveyor (MGS) orbiter [Malin et al., 1992], and the High-Resolution Stereo Camera (HRSC) on the Mars Express orbiter [Neukum and Jaumann, 2004] are most similar to THEMIS-VIS. The OMEGA imaging spectrometer on Mars Express [Bonello et al., 2004], the visible-band bolometer detectors of the Thermal Emission Spec-

trometer (TES) on MGS [Christensen et al., 1992], and the Wide-Field Planetary Camera 2 (WFPC2), Advanced Camera for Surveys (ACS), and Space Telescope Imaging Spectrometer (STIS) [e. g. Bell, 2003] on the Hubble Space Telescope (HST) are also currently providing (or have recently provided) visible-light coverage of Mars, but at significantly lower resolutions than MOC and HRSC.

MOC acquires monochrome images at scales as small as ∼1.5 m and two-color imagery at regional and global scales from a circular, 400 km altitude, 2:00 p.m. local solar time Sun-synchronous orbit [Malin et al., 1992]. HRSC provides stereo imagery and 4-channel color from a highly elliptical orbit with a best resolution of 50 meters per pixel [Neukum and Jaumann, 2004]. Both MOC and HRSC are line-scan cameras, which use linear Charge-Coupled Device (CCD) arrays to obtain information along one axis, and spacecraft ground track motion perpendicular to the CCD array direction to obtain the other spatial axis. THEMIS-VIS acquires monochrome or color images in any combination of its five bands with resolution modes of 18, 36, and 72 meters per pixel from its circular, 400 km, approximately Sun-synchronous orbit at typical local solar times of 4:00 p.m. to 5:00 p.m. The filter bandpasses have a FWHM of roughly 50 nm and are centered at 423, 540, 654, 749, and 860 nm. Fig. 4.1a shows the THEMIS-VIS bandpass profiles. Unlike MOC and HRSC, THEMIS-VIS is a 2-D array framing camera — all 1024 x 1032 CCD pixels are exposed simultaneously. However, spacecraft ground track motion is also used to expand the spatial coverage of acquired imaging sequences. While THEMIS-VIS is not designed to provide stereo capability, a limited amount of information on the altitude and cross-track velocity of high clouds can be derived by co-registering overlapping frames acquired at slightly changing viewing geometries over the course of a color imaging sequence (see Chapter 3).

THEMIS-VIS was found to have a serious problem from stray light contaminating the detector signal both during integration and during the CCD readout. Stray light accumulation during integration affects the entire array, but is most noticeable as broad and brighter stripes of additional signal near the edges of uncalibrated images. Stray light accumulation during the readout process creates an additive offset, and an effect similar to the "electronic shutter smear" seen in some shutterless CCD imaging systems, both of which are proportional to scene brightness but independent of the exposure time. Most of this chapter is focused on the removal of both forms of stray light contamination from THEMIS-VIS images.

This chapter describes and evaluates the process by which the radiance-calibrated data stored in the NASA Planetary Data System (PDS) THEMIS-VIS Reduced Data Records (RDRs; accessible via the internet at `http://themis-data.asu.edu/`) is generated from the raw PDS Engineering Data Records (EDRs). (Note that THEMIS-VIS RDRs are not geometrically calibrated. This chapter does not address geometric calibration.) We begin by describing the relevant operational details of THEMIS-VIS, as well as our preferred model for the stray-light contamination mechanisms. We then discuss each step of the EDR-to-RDR calibration process. Finally, we evaluate the calibration by comparing the derived THEMIS-VIS radiances to HST results, as well as to the surface-based multispectral measurements provided by the Mars Exploration Rovers' Panoramic Cameras (Pancams).

4.2 Operational Details and Labeling Conventions

VIS uses a Kodak interline-transfer CCD. For each column of photosites, there is a masked vertical register (v-register) adjacent to it. At the start of a THEMIS-VIS

exposure, any charge accumulated on the detector is flushed. Charge accumulated in the photosites is transferred to the adjacent v-registers twice during the commanded exposure period — once at the midpoint and once at the end. The v-registers are masked to minimize accumulation of photo-electrons during readout. During the readout process, beginning immediately after the second photosite charge transfer, the charge is shifted "downstream," *i.e.*, down the v-registers and transferred to the CCD's horizontal register (h-register) one row at a time. The h-register will be considered to be at the "bottom" of the detector array in this chapter, so that "down" and "downstream" are in the same direction. That is, if one row of pixels is said to be below or downstream of another row of pixels, then the row that is "below" or "downstream" is closer to the h-register. With this convention, the direction of spacecraft motion (south on the afternoon side of the orbit) is always "down." We number rows starting from zero at the bottom. Note, however, that the opposite row ordering convention is used in PDS EDR and RDR products.

Physically, the THEMIS-VIS CCD detector has 1032 columns and 1024 rows. The 5 THEMIS-VIS filters are strips that are 1032 columns wide and approximately 200 rows tall and are bonded directly to the detector. The front of the CCD housing is covered by a window, which therefore falls between the filter strips and the telescope optics, and which influences the overall spectral response in each THEMIS-VIS band. Figure 4.1a gives the spectral response through each THEMIS-VIS filter *with the spectral throughput of the CCD window included.* The spectral response of each filter by itself turns out to be important, as we will show, for predicting the magnitude of stray light contamination, and is therefore plotted in Figure 4.1b.

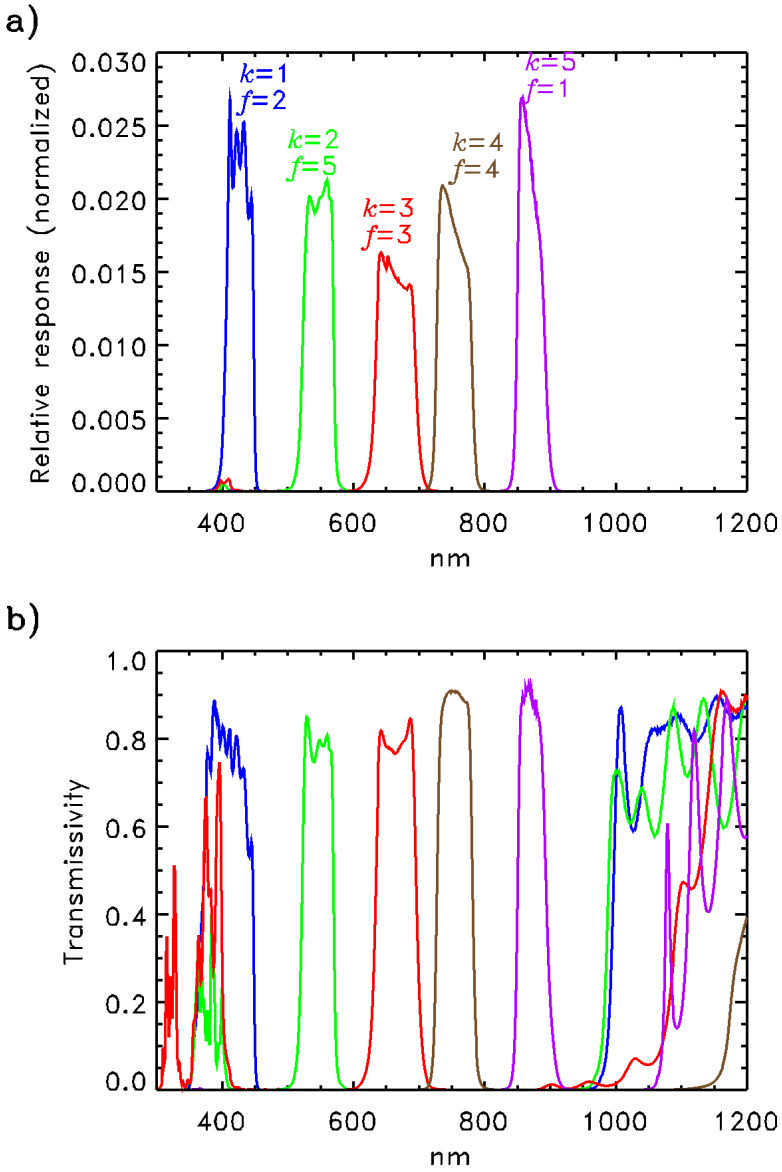

Figure 4.1: Relative response function (a), and filter transmissivity (b) for each band of the THEMIS-VIS system. Each response function curve is normalized so that the area under the curve equals one. The transmissivity curves represent the absolute transmission of each band's filter. The response functions and transmissivities are color coded, with blue, green, red, brown, and purple corresponding to the 425, 540, 654, 749, and 860 nm bands, respectively.

Figure 4.2: Photograph of the THEMIS-VIS detector and color filter array prior to THEMIS instrument assembly. Band (k) and filter (f) numbers are shown.

In this document, as in the EDR and RDR labels on the PDS files, the filters are labeled "filter 1", "filter 2", ... "filter 5" in order of their position on the detector array, with filter 1 at the bottom, closest to the h-register. In this naming scheme, "filter 1" is 860 nm, "filter 2" is 425 nm, "filter 3" is 654 nm, "filter 4" is 749 nm and "filter 5" is 540 nm. An alternate filter designation exists, using the labels "band 1", "band 2", ... "band 5" to refer to the filters in order of increasing center wavelength [425, 540, 654, 749, 860 nm]. Figure 4.2 shows the THEMIS-VIS focal plane with filters f and bands k labeled.

The groundtrack of the Odyssey spacecraft runs approximately parallel to the detector columns. This allows multispectral images to be acquired using a timed sequence of THEMIS-VIS exposures. Any combination of one or more filters can be acquired by selectively reading out only those rows that correspond to the desired filters. Thus, a single THEMIS-VIS exposure can contain any or all of the five

filters. We refer to the portion of an exposure that contains data for a single filter as a "framelet." All exposures in a sequence have the same exposure duration, and all filters in an exposure must have the same exposure duration, and so all framelets in a given THEMIS-VIS sequence have identical exposure durations.

The different filter framelets in a full-frame THEMIS-VIS exposure cover different regions of the Martian surface (nominally about 3.6 km apart for adjacent filters), hence the need for a sequence of multiple exposures to build up a multispectral image using spacecraft downtrack motion to image the same parts of the surface through different filters (and thus at slightly different times). The exposure sequences are designed so that all filters end up with the same number of framelets, and so that when a filter moves past the targeted region of the surface, it is no longer read out, meaning that the first and last several exposures of an imaging sequence will have a smaller number of filters read out than exposures in the middle of a sequence. Figure 4.3 illustrates the way that the framelets of a series of exposures map onto the Martian surface and can be arranged to form a continuous image strip for each filter.

The delay between exposures within a THEMIS-VIS imaging sequence is set so as to maximize the areal coverage of an imaging sequence while allowing for adequate overlap between adjacent framelets of the same band. A 1 second delay leads to a typical overlap of between 10 and 20 detector pixels (depending on the surface elevation and the angle of the planet's rotation relative to the groundtrack) when projected onto the martian surface. One side effect of this overlap is that framelets of different bands are not exactly spatially coincident when map projected, meaning that framelet boundaries (and calibration artifacts associated with those boundaries) for different bands are in different positions in mapped data, and

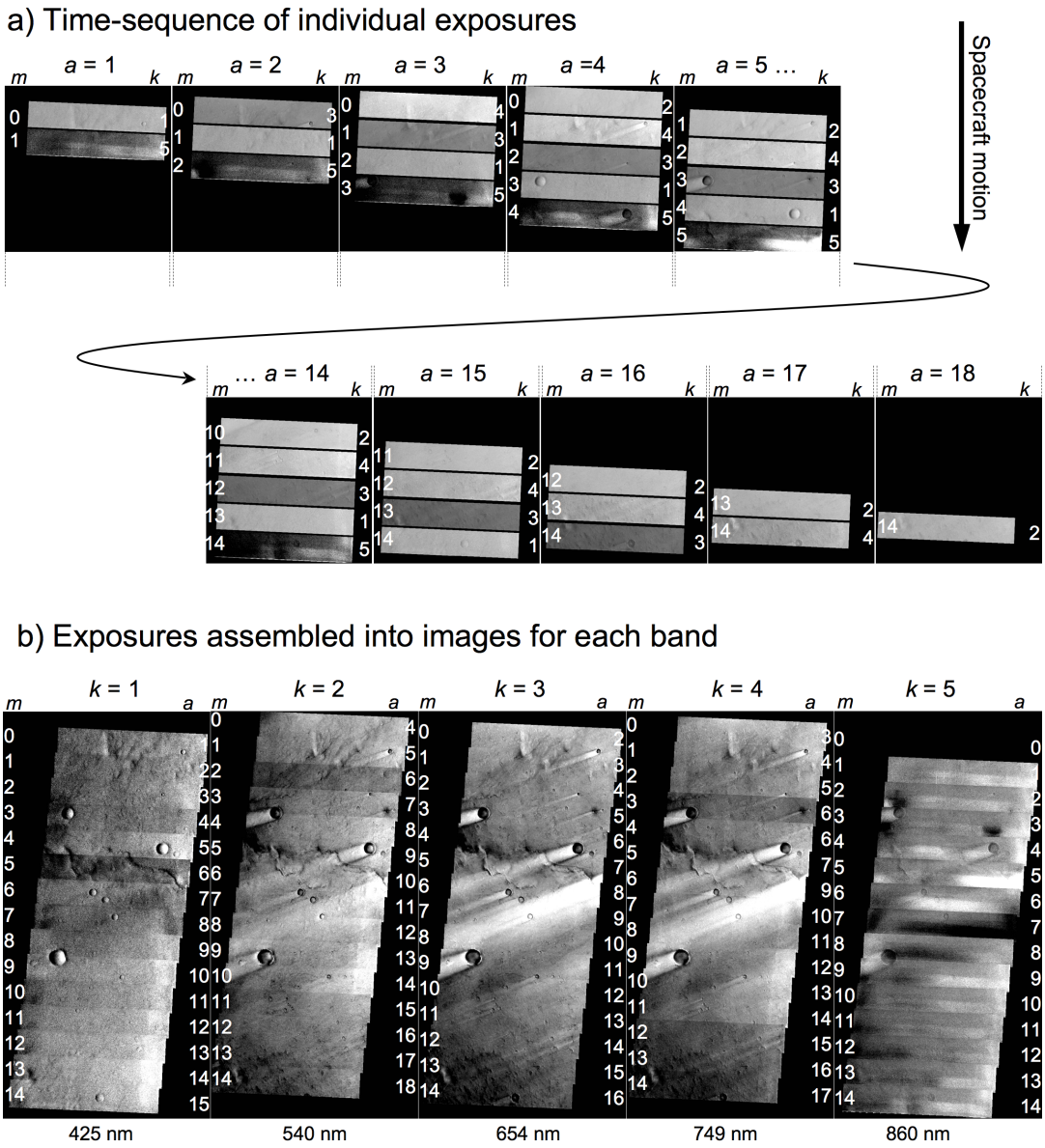

Figure 4.3: The individual exposures (a) that are used to construct THEMIS-VIS images for each band (b). Both the individual exposures and the assembled images for each band are THEMIS-VIS RDR framelets map projected onto a simple cylindrical coordinate surface, with north at the top and east at right. Each exposure in the time sequence is labeled with its exposure number, a. Each band is labeled with its band number, k. In the time sequence, each framelet is labeled with its framelet number, m, and its band number, k; and in the assembled images, each framelet is labeled with its framelet number, m, and its exposure number, a. For the individual exposures, the area shown for $a = 1$ to 5 corresponds to the top portion of the assembled images, and the area shown for $a = 14$ and up corresponds to the bottom portion of the assembled images. The $a = 0$ exposure is not shown because, as is often the case, $a = 0$ contains no valid data.

that the total area covered by lower-numbered filters within a sequence is offset relative to that covered by higher-numbered filters by a number of pixels equal to the framelet-to-framelet overlap multiplied by the difference in filter number. The delay between exposures is given by the "INTERFRAME_DELAY" keyword in the THEMIS-VIS PDS headers. A standard interframe delay of 1.00 seconds was for used image numbers up through V14885012. The standard interframe delay was changed to 0.90 seconds beginning with V14886001.

Rows and columns near the edges of the filters are always cropped during readout. The result is that each framelet has dimensions of 1024 by 192 detector pixels. The rows that are cropped are referred to henceforth as inter-framelet rows. When THEMIS-VIS EDR "cubes" are generated (3-dimensional data sets containing both spatial and spectral information, in PDS "QUB" file format), the THEMIS-VIS exposures in a sequence are broken up into their constituent framelets and re-assembled so that each framelet's position in the cube corresponds roughly to its spatial location on the surface of the planet. Thus, for a given filter, the corresponding plane of the EDR QUB is generated by concatenating the framelets from top to bottom in the order in which they were acquired. Assuming no spatial summing (spatial summing is discussed below), and using the PDS row ordering convention (opposite that used elsewhere in this chapter), the first (top) framelet goes in EDR rows 0 – 191, the second in EDR rows 192 – 383, etc.. Within a given EDR QUB plane, we assign the framelets a number, m, starting from the top, with the first framelet acquired in a given filter being $m = 0$.

The procedure for removing stray light accumulated during readout requires that all of the framelets from a given exposure be grouped together. Consider a framelet at position m_0 taken through filter number f_0. The position m_1 of the

framelet taken through filter number f_1 that was acquired in the same exposure as framelet (m_0, f_0) can be found by the following equation:

$$m_1 = m_0 + (f_0 - f_1)\,. \tag{4.1}$$

If $m_1 < 0$ or $m_1 \geq n$, where n is the number of framelets in each filter in the EDR cube, then there was no framelet of filter f_1 acquired in the same exposure as framelet (m_0, f_0). Note that equation 4.1 is written in terms of filter numbers, which are, as mentioned previously, not the same as band numbers. Note also, however, that the EDR and RDR QUB planes are stored in band order, not in filter order, and the location of each filter within the EDR and RDR cubes is given by the BAND_BIN_FILTER keyword in the PDS labels. Figure 4.4 shows all five planes of the EDR from a five-band THEMIS-VIS imaging sequence with each framelet labeled according to its band number k, filter number f, framelet number m, and exposure number a. Exposure number a is defined so that it is equal to m for the lowest numbered filter in a given EDR or RDR; thus:

$$a = m + \left(f - f_{\min}\right). \tag{4.2}$$

THEMIS-VIS can operate in any of three spatial summing modes. The preceding discussion has assumed summing mode 1, *i.e.*, no spatial summing, for clarity. In summing modes (*i.e.*, for spatial summing factors of) 2 and 4, the THEMIS-VIS pixels are binned 2 x 2 and 4 x 4, respectively. This creates framelets with 512 columns by 96 rows, or 256 columns by 48 rows, respectively. Summing mode 1 has a typical pixel scale of 18 meters on the martian surface for a nadir-pointed viewing geometry. Thus, summing modes 2 and 4 have nadir-pointed pixel scales of 36 and 72 meters per pixel. To date, the martian surface has been targeted solely in the nadir-pointed configuration, so actual pixel scales will vary only slightly from the

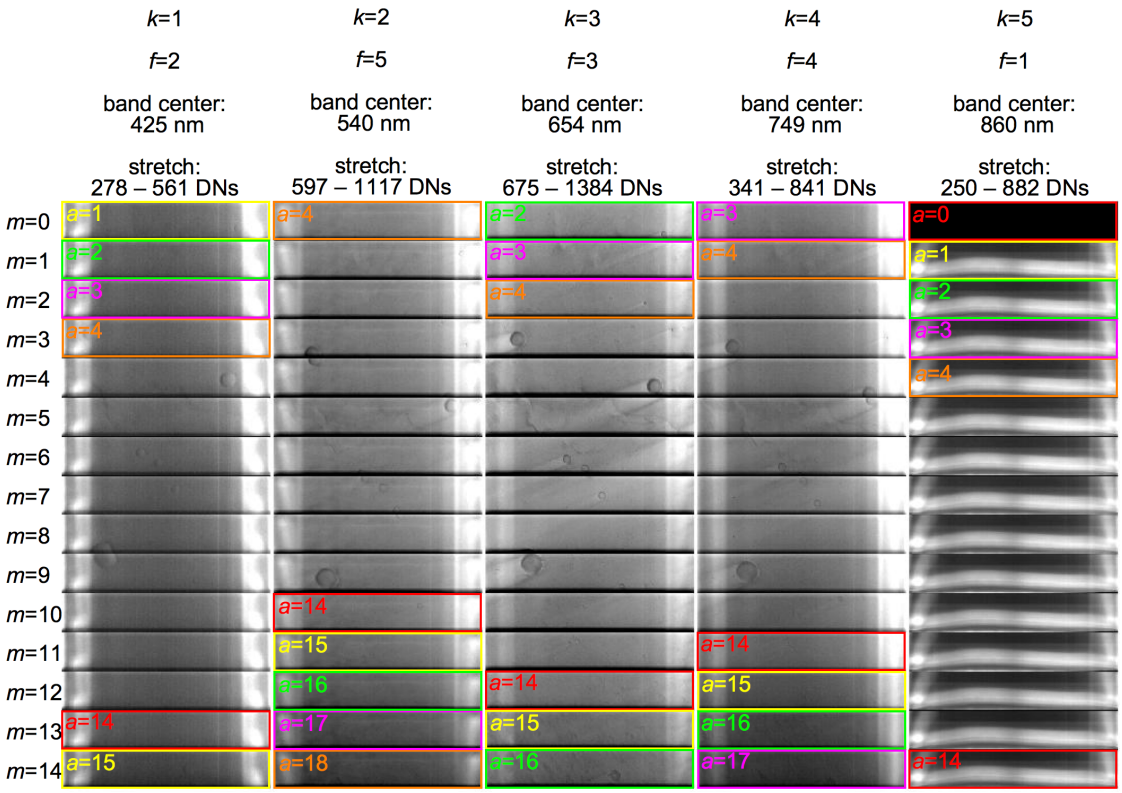

Figure 4.4: A THEMIS-VIS EDR for the same image as in Figure 4.3, with the five planes of the EDR layed out from left to right. To display this EDR data, we have performed the 8-bit to 11-bit conversion as described in the text and then applied a linear stretch to the DN values with the limits shown. We indicate filter number f and band number k for each band, and the colored boxes and labels indicate the exposure number a from which particular framelets were extracted.

values quoted here.

The amount of time required to read out a single THEMIS-VIS exposure depends both on the spatial summing mode and on the number of filters being read out. Each framelet of an exposure requires 141, 76, or 39 msec to read out, for summing modes 1, 2, and 4, respectively. A small amount of time ($\sim$4 msec) is also required to "dump" the charge from each filter that is not being read out. About 0.2 msec is required to dump each group of the inter-framelet rows, which are never read out. So, for example, it takes 703 msec to complete the readout process when summing mode 1 data from all five filters is being acquired. Since this readout time is long compared to the $\sim$5 – 10 msec typical exposure time, the opacity of the masks that protect the v-registers from photo-electrons, as well as the amount of time that each row of data spends in the v-registers during readout, are crucial factors for the THEMIS-VIS calibration process. And since THEMIS-VIS quickly dumps the charge in filters not being read out, the amount of time a given row spends in the v-registers depends not only on its distance from the h-register, but also on which of the filters that are below it on the array are being read out.

Correcting for stray light and dark current accumulation during readout requires knowledge of the timing for each data pixel as it is clocked down the v-registers to the h-register during readout. For pixels in a given filter, this timing depends on which of the lower-numbered (closer to h-register) filters are being read out in a particular exposure, and which are being dumped and not read out. Pixels in filter number f have $2^{((f-1)}$ possible timing "paths" during readout (*e.g.*, for filter 3, these are the read out filter combinations [3,2,1], [3,2], [3,1], and [3]). The

total number of possible paths is therefore

$$\sum_{f=1}^{5} 2^{(f-1)} = 31. \tag{4.3}$$

These 31 possible paths will henceforth be referred to as "filter paths." Each path is given a filter path code, F. For a given filter with filter number f_0, F is determined by:

$$\begin{aligned} F &= \sum_{f=1}^{f_0} \eta(f) \cdot 2^{(f-1)} \\ \eta(f) &= 1 \quad \text{IF filter } f \text{ is included in the exposure} \\ \eta(f) &= 0 \quad \text{IF filter } f \text{ is not included in the exposure.} \end{aligned} \tag{4.4}$$

A filter is included in an exposure if both of the following are true: (1) the filter is present in the EDR; and (2) the framelet position m of the filter (as calculated from equation 4.1) satisfies $m_1 < n$ and $m_1 \geq 0$, where n is the number of framelets of each filter in the EDR cube.

The calibration process is also influenced by the manner in which the THEMIS-VIS electronics perform spatial summing and store the data. The number of pixels summed in the "horizontal" direction (along rows) is always equal to the number of pixels summed in the "vertical" direction (along columns), but the spatial summing in each direction is performed in a different manner. Vertical summing, if any, is performed in the h-register. The charge from the appropriate number of v-register pixels is simply added together into each h-register site, and the h-register sites are then passed to the analog-to-digital (A-to-D) converter and represented as digital ("data number" or "DN") values with a gain measured at 25.4 electrons per DN. An additive bias equal to 104 DNs is removed at this stage, after which DN values can range from 0 to 2047 (*i.e.*, 11 available bits) regardless of the spatial summing mode. At this point, summing has only been performed along the vertical axis,

and the DN values represent charge from a number of pixels equal to the spatial summing factor. That is, for a given exposure time and scene brightness, 2x2 summing gives, neglecting charge accumulation during readout, twice as much h-register charge, and thus a DN value which is larger by a factor of two.

Next, horizontal summing is performed on these DN values by THEMIS-VIS onboard software, but now the appropriate number of pixels are averaged rather than summed, so that DN values in the 0 – 2047 range remain in the 0 – 2047 range. Thus, considering the total effect of the two disparate steps of the spatial summing process, the values returned by THEMIS-VIS with spatial summing are, with a given scene brightness and exposure time, still only larger by a factor equal to the spatial summing factor (rather than to the spatial summing factor squared) once readout-accumulated charge has been removed, and the amount of exposure time required to reach saturation of the available 0 – 2047 DN representation range is correspondingly less by a factor equal to the spatial summing factor.

Before the THEMIS-VIS data are sent to the Odyssey spacecraft flight computer for storage and eventual downlink, the 0 – 2047 11-bit DN values are compressed into a 0 – 255 8-bit range using a square-root encoding algorithm. It is important to note that 0 – 255 square-root encoded data is preserved in the EDR, that the encoding is non-linear, and that therefore the encoding must be reversed before the THEMIS-VIS EDR data is used in any quantitative way. (See Table 4.5.)

The number of framelets in any THEMIS-VIS imaging sequence is limited by the 3.8 megabyte capacity of the THEMIS-VIS instrument's data storage buffer, since THEMIS-VIS can transfer data to the spacecraft flight computer only after the completion of an imaging sequence. This maximum capacity corresponds to

19, 78, and 318 framelets for summing modes 1, 2, and 4 respectively, which means maxima of 3, 15, and 63 framelets per band for five-band imaging sequences, or, for example, 6, 26, and 106 framelets per band for three-band imaging sequences.

In addition to the normal operating mode described above, THEMIS-VIS has a "test" mode in which the entire 1024x1032 array is read out and stored, only a single exposure is acquired, and only spatial summing mode 1 is available. The in-flight test mode is otherwise identical to the normal operations. However, much of the calibration data used for this chapter was acquired using a pre-flight test mode with several additional differences from in-flight operations: 1) only one photosite transfer, at the end of the commanded exposure time, is performed; 2) no bias subtraction is applied prior to data storage; and 3) the data are never compressed to an 8-bit representation.

In order to compare pre-flight (or in-flight) test mode results with normal operating mode data, the test mode images are separated into framelets, with the test mode image yielding one framelet for each filter. These framelets are 1024x192, just like a normal framelet, and are obtained by extracting rows 2 – 193, 202 – 393, 403 – 594 , 611 – 802, and 813 – 1004, for filters 1 – 5, respectively, so as to cover the same region of the detector as the operating-mode framelets. Rows are numbered starting with zero at the bottom of the detector.

4.3 Stray Light Model

Our calibration method is based on the hypothesis that the source of the stray light signals is a light leak under the edges of the filters (Fig. 4.4). Since the edges of the filters are not masked, and since light reaches the focal plane at fairly high incidence angles (the THEMIS optics are an f/1.7 system), a significant portion

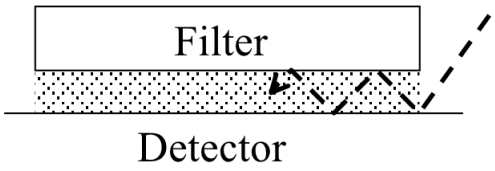

Figure 4.5: The hypothesized path followed by the stray light that reaches both the photosites and registers. The dashed line indicates one of many similar paths in which light reaching the edges of the detector is scattered under the edges of the filters. The diagram is not to scale.

of the light that strikes the focal plane near the edges of the filters is scattered or reflected towards the gap between the filter and detector (this gap is filled with a transparent adhesive used to bond the filters to the array; see Fig. 4.5). Since the interference filters strongly reflect broadband light, light that is scattered or reflected into the gap can be directed back down to the detector.

Light directed back towards the detector will obviously contaminate the photosites of the detector array, and we will refer to this effect as photosite stray light. Photosite stray light is illustrated in Fig. 4.6, which shows a pre-flight test mode image of a uniform integrating-sphere light source. Within the region covered by each filter the signal level is nearly uniform, and the brightness differences between the filters are the result of the calibration lamp spectrum and the wavelength dependence of the THEMIS-VIS system's responsivity. Superimposed on this expected pattern are broad brighter stripes at the bottom and on either side of the array. These brighter stripes are the result of the photosite stray light. Notice that the signal in most of the area of the 860 nm filter, located at the "bottom" of the array, is dominated by the photosite stray light. It is not known why there is no similar bright stripe in the 540 nm filter at the "top" of the array. Also note

that, as shown by Fig. 4.7, the amplitude of this stray light pattern is similar in each of the filters, despite the fact that the center-field, mostly uncontaminated photosite signal varies widely from filter to filter.

Light propagating in the filter-detector gap also contaminates the v-registers and h-register, since the narrow masks that protect the v-registers and h-register are less effective at high incidence angles. The v-register and h-register contamination is readily observed in the pre-flight calibration data by extrapolating DN vs. time to an exposure time of zero, thereby eliminating the photosite contribution to the signal. Fig. 4.8 shows a column profile of this register contamination signal for several different calibration-lamp brightnesses, with a 3-msec integration photosite signal shown for comparison. Fig. 4.9 shows an image of the register contamination signal for a particular pre-flight calibration lamp brightness. The ramp-up of the register contamination signal towards high row numbers seen in Figs. 4.8 and 4.9 is *superficially* similar to an electronic shutter smear effect, but it differs in that it depends on the brightness of the scene just outside of the detector field of view, rather than on the portion of the scene imaged by the detector, and so its magnitude and behavior are not directly related to the information contained in the image.

Although details of the scattering/reflection processes and the properties of the register masks are not well known and so must be derived as part of the calibration procedure, the general features of our model yield several key assumptions for that procedure:

1. The source for both types of stray light is just outside the field of view of the detector. Therefore, the calibration procedure can never exactly remove stray light effects because the scene radiance in the region just outside the

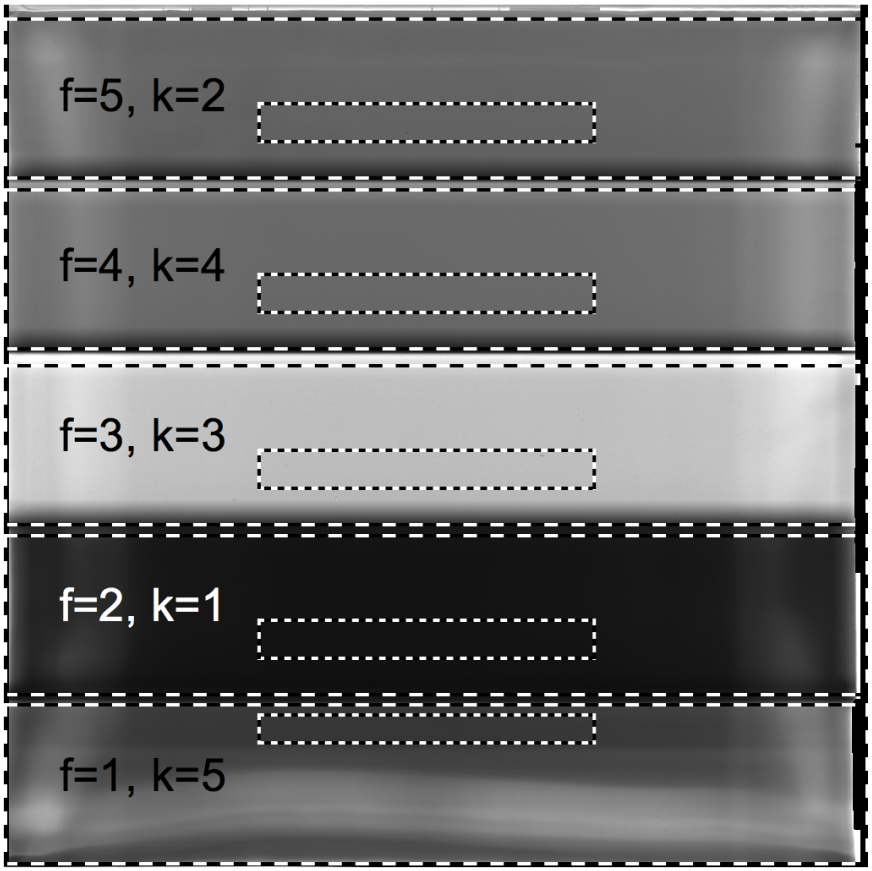

Figure 4.6: Pre-flight test-mode image showing the photosite signal from the integrating sphere calibration source (using nine 8 W lamps), with each filter labeled with its filter (f) and band (k) number. The large boxes outline the regions of the detector from which the 1024 $\times$ 192 framelets for each filter are extracted. The smaller boxes show the boundaries of the "calibration region of interest" (C-ROI) for each filter. The inter-filter differences in brightness are due to the spectrum of the light source and the transmissivities of the filters. However, all of the obvious intra-filter non-uniformity in this image is attributable to photosite stray light. This photosite stray light forms the broad stripes on the left and right side of all of the filters as seen in this image. It also forms the bright stripe that covers the bottom half of the $f = 1$ filter.

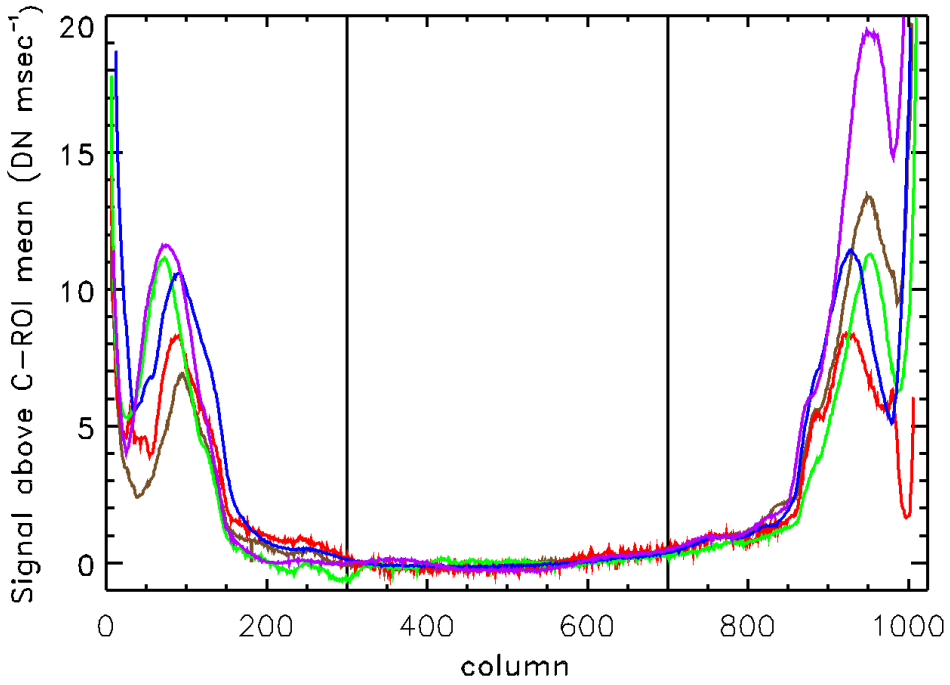

Figure 4.7: Photosite signal as a function of column in the pre-flight test mode image in Figure 4.6. The signal for each band is color coded, with blue, green, red, brown, and purple corresponding to $k = 1$, 2, 3, 4, and 5 (425, 540, 654, 749, and 860 nm), respectively. The column signal is averaged over the rows included in each band's C-ROI, and is shown as the difference between the column signal and the given band's C-ROI mean. The vertical lines indicate the boundaries of the C-ROIs (which are identical in the column dimension for all 5 bands).

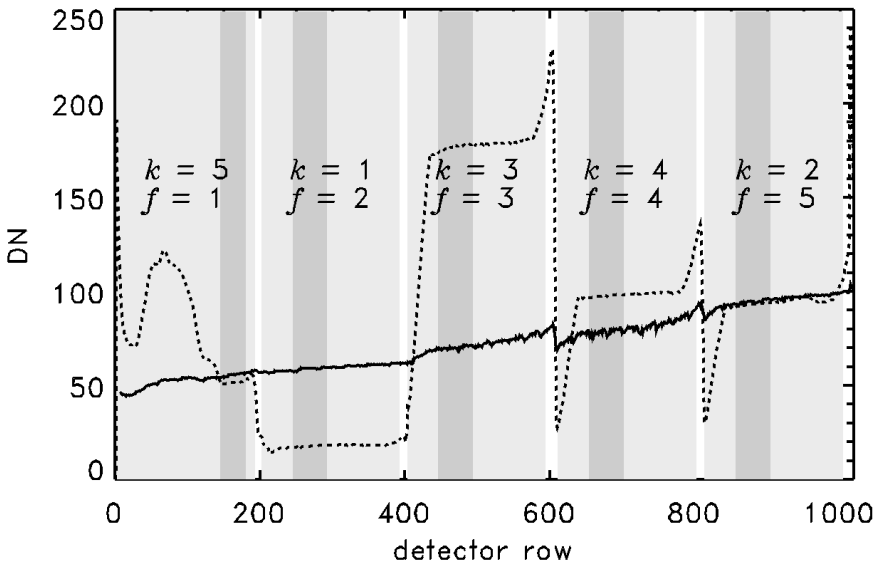

Figure 4.8: Register signal (solid lines) and photosite signal (dashed lines) as a function of row number, for a 3 msec exposure using pre-flight test-mode and nine 8 W lamps. The signal is averaged over all columns included in the C-ROIs. The shaded regions indicate the rows from which the framelets are extracted, and the darker shaded regions indicate the rows included in each C-ROI. Each shaded region is labeled with its filter (f) and band (k) numbers.

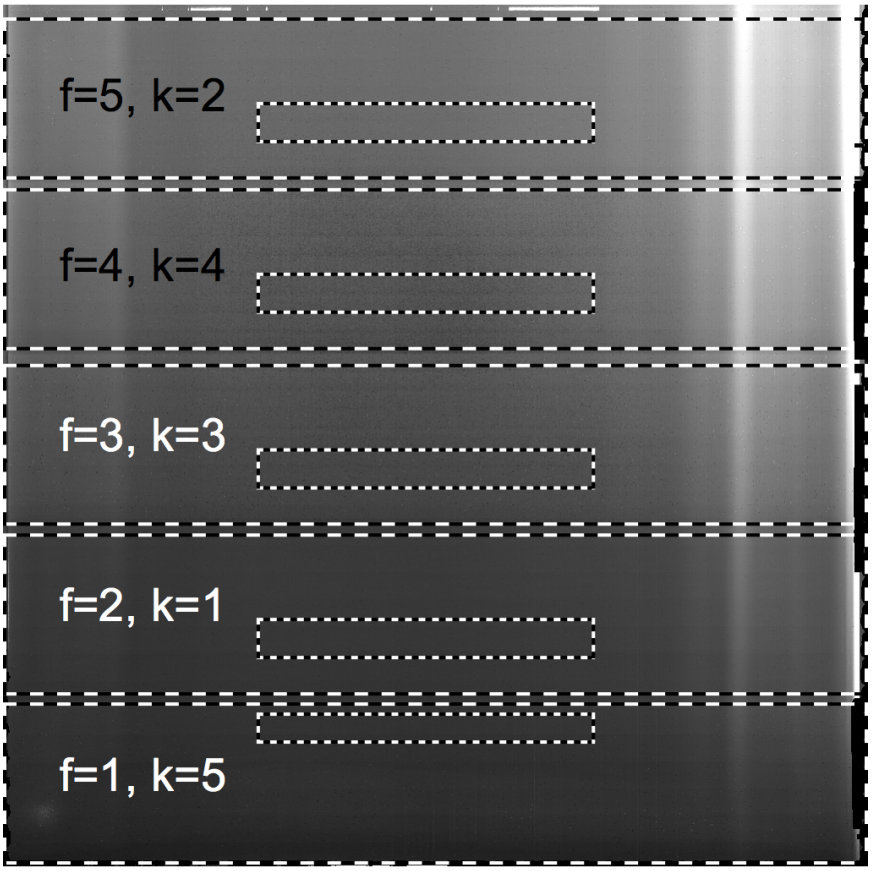

Figure 4.9: Pre-flight test-mode image showing the register signal from the integrating sphere calibration source (using nine 8 W lamps), with each filter labeled with its filter (f) and band (k) number. The large boxes outline the regions of the detector from which the 1024 $\times$ 192 framelets for each filter are extracted. The smaller boxes show the boundaries of the C-ROI for each filter.

field of view can by definition only be estimated. For low-contrast scenes, this "guess" can be fairly accurate, but for high-contrast scenes, the error inherent in estimating out-of-field stray light can be significant.

2. The register stray light is proportional to the scene radiance, and to the amount of time that each row spends in the v-registers during the readout process, but it does not depend on the exposure time (since the readout time depends only on the spatial summing mode and filter combination).

3. The photosite stray light is proportional to both the scene radiance and the exposure time.

4. Both stray light signals are proportional to the "broadband" scene radiance, rather than the narrow-band radiance in any of the five filters, since the stray light bypasses the filters in our model of the effect. Thus, just as the "direct" (*i.e.*, stray-light free) signal in each filter is proportional to the scene radiance spectrum weighted by that particular filter's response function (shown in Fig. 4.1a), the stray light signal in each filter is proportional to the scene radiance weighted by a different response function — one which includes all of the same factors as those for the filters, except for the transmissivity of the filters themselves. This response function is shown in Fig. 4.10. The dominant factors in this broadband response function are the CCD window and the detector quantum efficiency.

4.4 Calibration Procedure

The THEMIS-VIS flight-data calibration process, (henceforth, the "calibration pipeline") consists of seven steps: (1) 8-bit to 11-bit decoding; (2) identification of

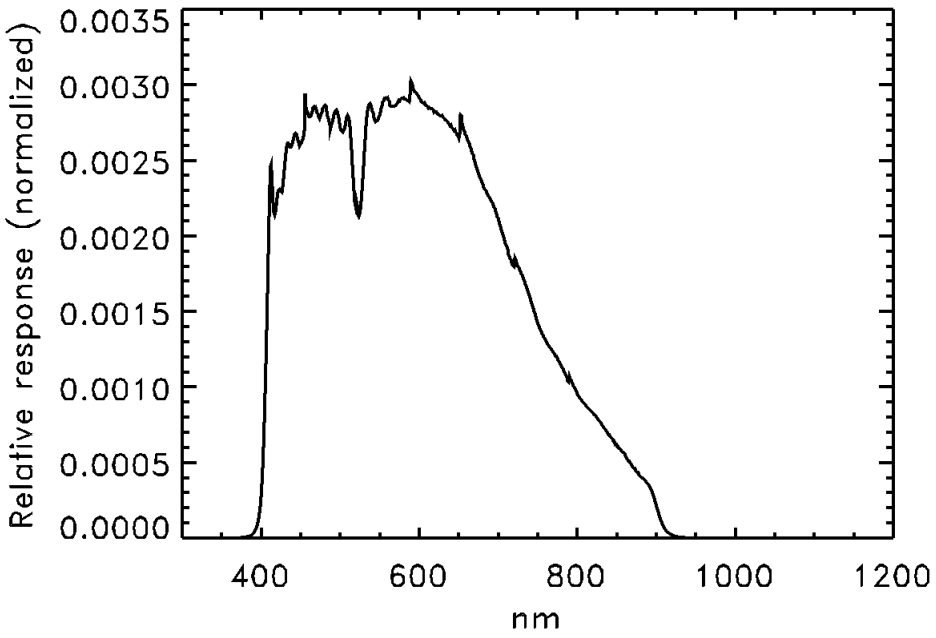

Figure 4.10: Relative response function of the THEMIS-VIS system excluding the filter transmissivities, *i.e.*, the broadband response function. The response function is normalized so that the area under the curve equals one.

bad pixels; (3) bias subtraction; (4) register stray light subtraction; (5) correction for pixel response variations (*i.e.*, "flatfielding"); (6) photosite stray light subtraction; and (7) conversion to radiance. Derivation of the calibration frames and calibration coefficients necessary for this pipeline proceeds in a somewhat different order, however. The first step of the calibration derivation is to use the pre-flight data set to determine the photosite stray-light response and the direct response for the central region of each filter. The next step is to develop a model, which we will call the "broadband radiance model" that predicts the broadband scene radiance from the narrow-band scene radiances in each of the five THEMIS-VIS bandpasses. The response coefficients and broadband radiance model are necessary for deriving the calibration frames of pipeline steps 4, 6, and 7. Once the response coefficients and broadband radiance model are in hand, the calibration frames are derived in the same order as the pipeline steps using only on-orbit data, and so we will conclude this section by describing each pipeline step together with the derivation of any required calibration frames.

Many of the calibration coefficients, and in particular the response coefficients, apply to the entire area of a filter, and so they are calculated using mean values from a representative sub-region of that filter's framelets. This sub-region should also contain a minimal amount of stray light, so that the calibration coefficients based on it are influenced as little possible by any uncertainties in the stray light estimates. We have identified one such representative-yet-minimal-stray-light region, which will be referred to as the "calibration region of interest" (C-ROI), for each filter. The C-ROI boundaries are shown in Fig. 4.6.

4.4.1 Derivation of response coefficients from pre-flight data

Given the differences between the pre-flight and on-orbit operating modes, and considering the possibility of launch or post-launch changes in instrument behavior, we have derived the THEMIS-VIS calibration using on-orbit data wherever possible. However, in order to accurately determine the photosite radiometric responses, we require a source with known spectral radiance. In principle, this requirement could be met on-orbit using radiance measurements acquired by another instrument, but this other instrument would need to observe the same region of Mars at the same time as a THEMIS-VIS imaging sequence, with the same viewing geometry, and have a well validated calibration, high spectral resolution, and sub-kilometer scale spatial resolution. We have acquired concurrent THEMIS-VIS – WFPC2 and THEMIS-VIS – ACS measurements that come closest to satisfying these criteria, but given the spatial resolution discrepancy, and since the difference in phase angle between the HST measurements and the THEMIS-VIS measurements is very large, and the high-phase photometric properties of the martian surface-atmosphere system are potentially variable and not well known, we deem our ground based spectral radiance source to be the most reliable means for determining the radiometric response. The HST data sets will be used to validate the results.

The pre-flight data set we have used was acquired with the THEMIS instrument in a temperature-controlled vacuum chamber viewing an integrating-sphere light source through a window in the chamber wall. The integrating-sphere exit port filled the THEMIS-VIS field of view and extended well beyond it in all directions. We verified that changing the distance between the integrating sphere exit port

and the THEMIS-VIS window had no effect on the observed signal pattern or signal magnitude.

The integrating sphere output spectral radiance is varied by changing the number of 8, 45, and 200 Watt lamps, and for many of the possible lamp combinations the integrating sphere output is traceable to National Institute of Standards and Technology (NIST) standards. For the THEMIS-VIS calibration activities, we used from one to nine 8W lamps, and also settings with one 45W lamp and two 45W lamps. The lamp settings with the 8W lamps are all directly NIST-traceable with $\pm 2\%$ precision. The one 45 W and two 45 W settings are *not* directly NIST-traceable. We have estimated the spectral radiance for these settings by subtracting the NIST-traceable ten 8W setting spectral radiance from the NIST-traceable ten 8 W, one 45 W setting and ten 8 W, two 45 W setting spectral radiances. This subtraction is of course only valid under the assumption that the lamp settings are linear combinations of each other. We have tested the linearity assumption by differencing the NIST-traceable radiance measurements of sets of 8W lamp settings; for example, nine 8 W $-$ six 8 W $-$ three 8 W should give zero radiance with a precision degraded by a factor of roughly $\sqrt{3}$. We find that the linearity assumption is valid under such tests, and we therefore conclude that our differencing-based spectral radiance values for the 45 W lamps will have a precision of roughly $\sqrt{2} \times \pm 2\%$, *i.e.*, better than $\pm 3\%$, and will not be systematically biased. In any case, the 45 W lamps provide a different spectral shape from the 8 W lamps, and this different spectral shape is crucial for separating the direct response from the stray-light response, so we have no choice but to use the 45 W lamps in the calibration process.

The spectral radiance presented to THEMIS-VIS is also of course affected by

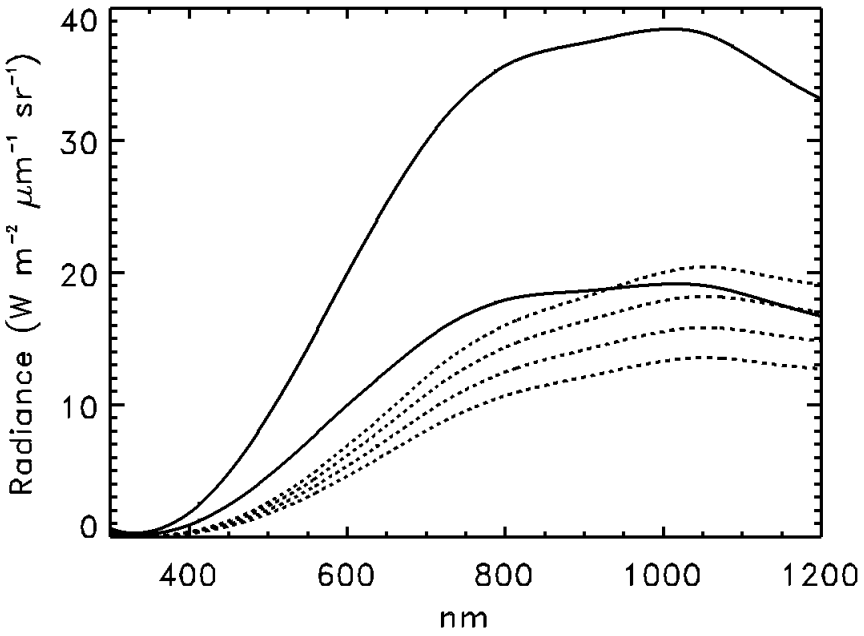

Figure 4.11: Spectral radiance of the integrating sphere calibration source for (bottom to top) six 8 W lamps, seven 8 W lamps, eight 8 W lamps, nine 8 W lamps, one 45 W lamp, and two 45 W lamps. Solid lines are used for the 45 W lamps, and dashed lines for the 8 W lamps.

the transmissivity of the vacuum chamber window. Fig. 4.11 shows the integrating sphere radiance for each lamp setting after correction for the window transmissivity. For each lamp setting, we calculate a weighted mean broadband radiance, as well as a weighted mean narrow-band radiance for each filter, by weighting the window-corrected sphere radiances with the response functions of Fig. 4.11 and Fig. 4.11, respectively. These weighted mean radiance values are listed in Table 4.1.

A set of THEMIS-VIS integrating-sphere tests was performed in a thermally controlled vacuum chamber temperature with the THEMIS-VIS focal plane temperature maintained at each of six different settings. For each temperature setting, measurements were taken with the following twelve lamp settings: all lamps turned

Table 4.1: Radiance Calibration Data Set

	Band					
	425 nm	540 nm	654 nm	749 nm	860 nm	
Lamp configuration	In-band radiance ($\mathrm{W\,m^{-2}\,\mu m^{-1}\,sr^{-1}}$)					Broadband radiance ($\mathrm{W\,m^{-2}\,\mu m^{-1}\,sr^{-1}}$)
6 8W	0.517	2.947	6.57	9.595	11.706	4.845
7 8W	0.599	3.434	7.777	11.212	13.692	5.650
8 8W	0.690	3.941	8.926	12.885	15.779	6.492
9 8W	0.770	4.411	9.985	14.424	17.544	7.258
1 45W	1.624	7.008	12.996	16.833	18.468	9.559
2 45W	3.259	13.917	26.087	33.486	36.958	19.088
	Signal ($\mathrm{DN\,msec^{-1}}$); 250 K thermal-vac test					
6 8W	3.612	19.436	37.551	20.782	10.757	
7 8W	4.236	22.645	43.740	24.152	12.585	
8 8W	4.799	25.646	49.273	27.258	14.204	
9 8W	5.267	28.437	55.645	30.277	15.761	
1 45W	9.626	44.638	73.635	36.276	19.444	
2 45W	19.150	89.408	146.314	72.193	38.787	
	Signal ($\mathrm{DN\,msec^{-1}}$); 268 K thermal-vac test					
6 8W	3.779	19.919	39.090	22.010	11.486	
7 8W	-	-	-	-	-	
8 8W	4.863	25.960	50.684	28.649	15.013	
9 8W	-	-	-	-	-	
1 45W	9.683	44.913	75.706	37.823	20.344	
2 45W	19.406	90.219	150.652	75.583	40.937	
	Signal ($\mathrm{DN\,msec^{-1}}$); 279 K thermal-vac test					
6 8W	3.652	19.664	39.113	22.399	11.637	
7 8W	4.246	22.958	45.592	26.042	13.556	
8 8W	4.778	25.901	50.949	29.456	15.384	
9 8W	5.365	28.855	58.287	32.719	17.084	
1 45W	9.663	45.316	77.247	39.327	21.024	
2 45W	19.340	90.951	153.461	78.483	42.327	
	Signal ($\mathrm{DN\,msec^{-1}}$); 295 K thermal-vac test					
6 8W	3.634	19.649	39.728	23.084	12.076	
7 8W	4.367	23.423	46.920	27.351	14.372	
8 8W	4.940	26.518	51.224	30.959	16.252	
9 8W	-	-	-	-	-	
1 45W	9.728	45.605	79.233	41.058	22.101	
2 45W	19.375	98.835	157.206	81.664	44.265	

off, one through nine 8W lamps, one 45W lamp, and two 45W lamps. Exposure times of 3, 6, 12, and 24 milliseconds were used at each lamp setting, and ten images were acquired at each exposure time. All ten images are averaged together to produce a single low-noise image at each exposure time. We use this data set to estimate the photosite stray-light response and the direct response as follows:

1. Fit a line to the set of exposure times for each pixel of each lamp level in order to determine S^*, the total photosite (stray-light plus direct) signal, in DN per msec.

$$D(t)_{ijk,l} = Z_{ijk,l} + S^*_{ijk,l} \cdot t + \epsilon_{ijk,l} \,, \tag{4.5}$$

   where $D(t)_{ijk,l}$ is the observed raw DN value for pixel i, j of band k at lamp level l, t is exposure time, and the residuals of the fit are represented by ϵ. Z gives the register stray-light signal. Z is useful for understanding the behavior of the register stray light, but since the pre-flight test-mode readout pattern and timing is different from the flight operating mode, Z is not directly applicable to the flight data calibration.

   Since the pre-flight test mode performs only one photosite-to-register transfer in the course of an exposure, detector full-well was reached well before the A-to-D conversion saturation value of 2047. We observed non-linear response associated with full-well at DN levels greater than 1250. Therefore, we inspected the $D(t)_{ij}$ images for each filter and lamp level and excluded any with substantial regions of > 1250 DN. Additionally, in the linear fitting process, we exclude all individual pixels that are either greater than 1250 DN, or which have more than 5% of pixels > 1250 in the surrounding 11x11 region.

2. Determine the background-subtracted total photosite signal, S:

$$S_{ijk,l} = S^*_{ijk,l} - S^*_{ijk,0} \,, \tag{4.6}$$

where $S^*_{ijk,0}$ refers to the lamp setting with all lamps turned off. The zero lamp signal is clearly not dark current, as the zero lamp images show a stray light pattern similar to the directly illuminated images, and their peak center-filter values of 2.5 DN per msec at 262 K are too high for plausible dark current and shows no temperature dependence. The zero lamp signal is also spectrally distinct from the integrating-sphere signal, which implies that its source may be background light in the testing room. In addition to subtracting this background light, we have also attempted to mitigate its impact by excluding the lower lamp settings from the model fits described below. Thus, in deriving the response coefficients, we use only those settings with the 45W lamps or at least six 8W lamps. With this restriction, the largest background signal contribution is to the 423 nm filter signal at six 8W lamps, where it makes up ∼15% of the total signal. Although the background subtraction should remove most of this contribution, without knowing the source of the background signal we have no way to assess its variability aside from the residuals of the model fits described below.

3. For each band k, extract the C-ROI mean values of the background subtracted photosite signal, $\overline{S}_{k,l}$ ("Signal" in Table 4.1). Then fit them to a linear model, with weighted-mean broadband radiance $\widehat{\overline{I}}_l$ (in Table 4.1), and weighted-mean band k narrow-band radiance $\overline{I}_{k,l}$ (also in Table 4.1) as the independent variables:

$$\overline{S}_{k,j} = x_k \, \widehat{\overline{I}}_l + y_k \overline{I}_{k,l} + \epsilon_{k,l} \,, \tag{4.7}$$

where x_k and y_k are the photosite stray light response coefficient and direct response coefficient, respectively. The straight bar in the symbols $\widehat{\overline{I}}_l$ and $\overline{I}_{k,l}$ indicates that these radiances should be interpreted as spatial means over the region sampled by a framelet.

Since $\widehat{\overline{I}}_l$ and $\overline{I}_{k,l}$ are highly correlated, we have taken the precaution of solving for x_k and y_k using a grid-search algorithm, which allows us to map the solution set and visualize the extent to which x_k and y_k can be simultaneously constrained. We perform the grid-search with a grid spacing of 0.005 $\mathrm{DN\,msec^{-1}/\left(W\,m^{-2}\,\mu m^{-1}\,sr^{-1}\right)}$ in both x_k and y_k, calculating the χ^2 probability at each point. The search range was 0 – 3 $\mathrm{DN\,msec^{-1}/\left(W\,m^{-2}\,\mu m^{-1}\,sr^{-1}\right)}$ in x_k and 0 – 7 $\mathrm{DN\,msec^{-1}/\left(W\,m^{-2}\,\mu m^{-1}\,sr^{-1}\right)}$ in y_k. In order to calculate the χ^2 probability, each data point is assigned a standard error equal to the root-mean-squared residual of the best fit. Fig. 4.12 shows the resulting normalized probability density function (PDF) for 4 different vacuum chamber temperatures. Note that the model fits do a poor job of constraining the photosite stray-light response for bands 2 and 4, and that they provide no constraint at all on the photosite stray light response of band 3. Also note the small upward trend of direct response with increasing temperature for bands 2 through 5. Since the 268 K and 279 K tests bracket the observed on-orbit focal plane temperatures of 268 – 278 K, we will use these tests to determine our response coefficients and associated confidence intervals.

4. Identify the band 1 direct and stray light response coefficients directly from the 279 K PDFs. Since band 1 shows no temperature trend for either coefficient, we are free to choose the most precise estimate, which is the one

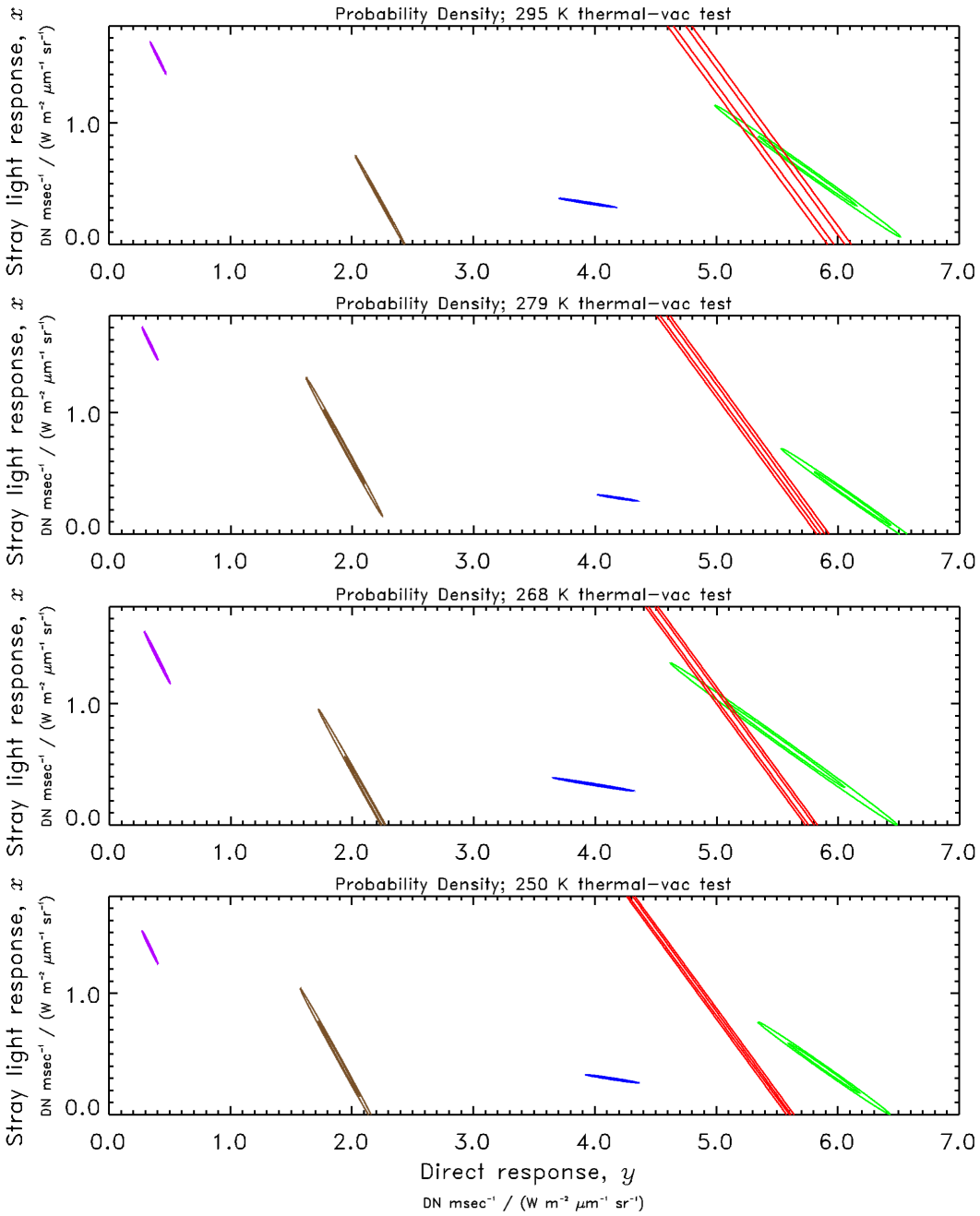

Figure 4.12: Two-dimensional probability density functions for the response coefficient fits at four thermal-vac temperatures. The PDFs are color coded, with blue, green, red, brown, and purple corresponding to $k = 1$, 2, 3, 4, and 5 (425, 540, 654, 749, and 860 nm), respectively. The contour lines represent χ^2 probabilities of 0.05, 0.30, 0.60, and 0.90.

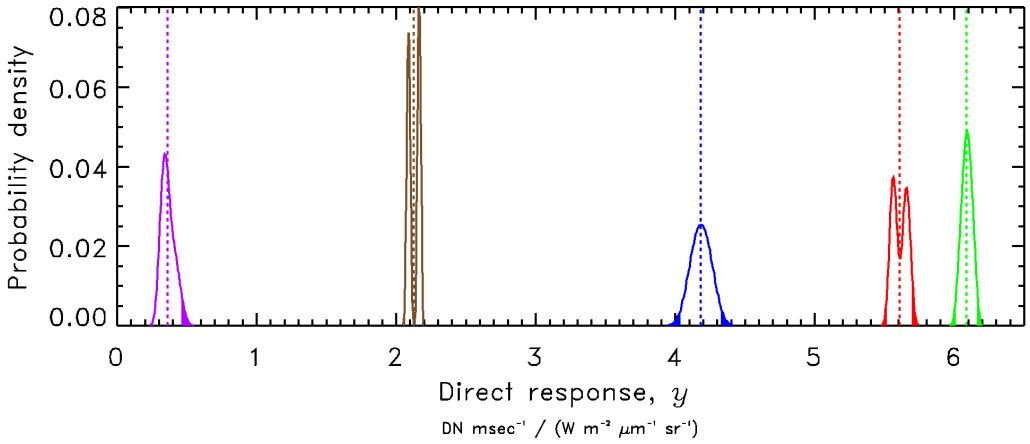

Figure 4.13: One-dimensional probability density functions for yk, the direct response coefficients. The PDFs are color coded, with blue, green, red, brown, and purple corresponding to k = 1, 2, 3, 4, and 5 (425, 540, 654, 749, and 860 nm), respectively. Dashed lines show the adopted y values, which correspond to the mean of each PDF. The adopted confidence intervals for y encompass 95% of the total probability and are centered on the PDF mean. The filled regions of each PDF represent regions outside of the adopted confidence intervals.

calculated at 279 K. The normalized 1-D PDF for direct response is shown in Fig. 4.13, and the normalized 1-D PDF for stray light response in Fig. 4.14. The adopted values for the coefficients, and their confidence intervals, are given in Table 4.2. The adopted values are the mean of the population described by the 1-D PDF. The confidence intervals are ranges centered on the adopted value and encompassing 95% of the total probability in the 1-D PDF.

5. Identify the band 5 response coefficients by combining the 1-D PDFs for the 268 K and 279 K tests. This is necessary because the band 5 coefficients exhibit a temperature trend. The adopted values and confidence intervals are derived as previously described, but using the sum of the normalized 1-D PDFs for the two tests. However, as will be described later, the band 5

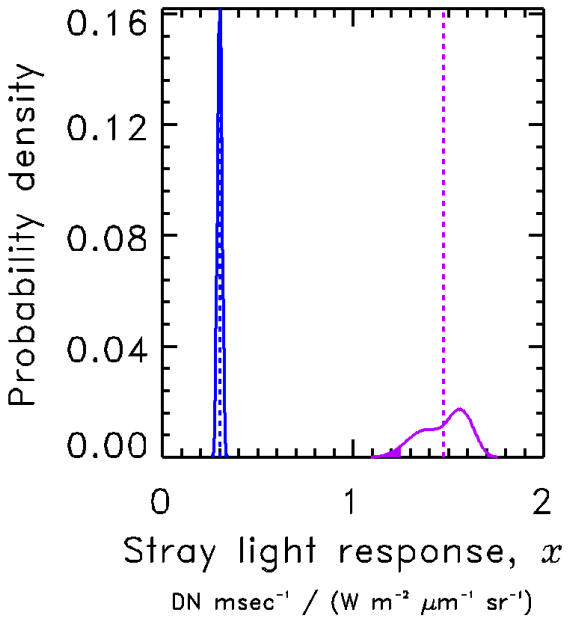

Figure 4.14: One-dimensional probability density functions for x, the stray light response coefficients. The PDFs are color coded, with blue corresponding to $k = 1$, and purple corresponding to $k = 5$. Dashed lines show the adopted x values, which correspond to the mean of each PDF. The adopted confidence intervals for x encompass 95% of the total probability and are centered on the PDF mean. The filled regions of each PDF represent regions outside of the adopted confidence intervals.

direct response coefficient derived in this manner was found to conflict with on-orbit validation measurements. We have therefore chosen to adopt the band 5 direct response coefficient implied by the validation measurements.

6. Assume that band 2 – 4 have the same stray light response as band 1. Some kind of assumption about the band 2 –4 stray light response is obviously required, since it is poorly constrained by the integrating sphere data set, and we have chosen the simplest possible assumption that is consistent with the 2-D PDFs. The row profiles of bands 1 – 4 shown in Fig. 4.7 indicate that these four bands show similar stray light patterns with similar amplitudes, which suggests that similar amounts of light are reaching the central area of each filter, and that our assumption is therefore reasonable.

7. Identify the direct response coefficients for bands 2 – 4 by combining the 1-D PDFs for the 268 K and 279 K tests. For these bands, the 1-D PDF is generated by summing over the region of the 2-D PDF that falls within the band 1 stray light response coefficient confidence interval.

Table 4.2: Adopted Response Coefficients

	Band				
	425 nm	540 nm	654 nm	749 nm	860 nm
x	4.180±0.145 [1]	6.085±0.075	5.605±0.090	2.125±0.060	0.6±0.2
y	0.300±0.025	0.300±0.025	0.300±0.025	0.300±0.025	1.475±0.225
	units: $\left(\mathrm{DN\,msec^{-1}}\right)$ / $\left(\mathrm{W\,m^{-2}\,\mu m^{-1}\,sr^{-1}}\right)$				

	Summing Mode		
	1	2	4
z	5.50±0.48	6.70±0.44	8.40±0.99
	units: DN/ $\left(\mathrm{W\,m^{-2}\,\mu m^{-1}\,sr^{-1}}\right)$		

[1] The given ranges represent 95% confidence intervals.

4.4.2 Derivation of broadband radiance model from HST data

Since the broadband radiance, $\widehat{\overline{I}}$, is not generally known during on-orbit operations, it must be estimated from the narrow band radiances, $\overline{I}_k$. We do this by finding a set of coefficients α_k such that

$$\widehat{\overline{I}}_q = \sum_k \alpha_k \overline{I}_{k,q} + \epsilon_q \, . \tag{4.8}$$

Alternatively, we can estimate $\widehat{\overline{I}}$ directly from the photosite signal, S:

$$\widehat{\overline{I}}_q = \sum_k \omega_k \overline{S}_{k,q} + \epsilon_q \, . \tag{4.9}$$

The ω_k coefficients are related to the α_k coefficients by

$$\omega_k = \frac{(\alpha_k / y_k)}{1 + \sum_k x_k \, (\alpha_k / y_k)} \, , \tag{4.10}$$

which can be shown by combining Equations (4.7), (4.8), and (4.9).

The α_k or ω_k can be estimated from a data set of $\overline{I}_k$ and $\widehat{\overline{I}}$ that is representative of the radiance of Mars during on-orbit operations. We have chosen a set of HST WFPC2 and ACS observations of Mars performed in 2003 during the Odyssey mission [Bell et al., 2004b], the details of which are given in Table 4.3. These HST observations use narrow-band filters to sample the Mars radiance spectrum, and so we must interpolate from these samples to form a complete spectrum suitable for convolution with the THEMIS-VIS narrow-band and broadband spectral response functions. We do this by performing a cubic spline interpolation of the HST I/F samples and then multiplying by the solar radiance spectrum.

The HST measurements are calibrated using the methods described in Bell et al. [1997]. The images are projected onto a 2 pixel-per-degree simple cylindrical

grid, and the Mars radiance spectrum at each grid point is used to generate $\overline{I}_k$ and $\widehat{\overline{I}}$ for each grid point. All grid points from all of the HST observation sequences described in Table 4.3 form the $\widehat{\overline{I}}_q$ and $\overline{I}_{k,q}$ for the multiple linear regression model indicated by Equations (4.8) or (4.9). We have used this regression model approach to find α_k and ω_k in two ways: 1) using Eq. (4.7) to calculate $\overline{S}_{k,q}$ from $\widehat{\overline{I}}_q$ and $\overline{I}_{k,q}$, then Eq. (4.9) to solve for ω_k, and then Eq. (4.10) to give α_k from ω_k; and 2) using Eq. (4.8) to solve for α_k and then Eq. (4.10) to give ω_k. Both methods produce essentially identical results.

Our adopted ω_k coefficients are shown in Table 4.4. In order to produce the best possible estimate of $\widehat{\overline{I}}$ for a particular image sequence, the calibration pipeline needs to be able to use whatever $\overline{S}_k$ values are available, and so we need a different set of ω_k for each of the 31 possible band combinations. The calibration pipeline will choose the set of coefficients for the band combination that includes only those bands that are being used to estimate $\widehat{\overline{I}}$, regardless of which bands happen to be present in an EDR or an exposure. Since the ω_k are regression coefficients and *not* response coefficients, negative values are in principle allowed. However, all of the negative coefficients in Table 4.4 are very small relative to the others for their band combination, and so can be interpreted as zero, meaning that a particular band when part of a particular band combination does not contribute significantly to the $\widehat{\overline{I}}$ estimate.

4.4.3 Flight data calibration procedure

In this subsection, we describe the calibration pipeline procedures in the order that they are performed.

Table 4.3: 1997 and 2003 HST WFPC2/PC and ACS/HRC Observations of Mars Used in this Study

UT Date[1]	UT Time[2]	Camera	Filter Wavelengths nm	Diam. arc-sec.	SE Lat. (°)	SE Lon. (°)	Phase Angle (°)	L_s (°)	Res.[3] km/pix	PROG ID[4]
970310	6:43	WFPC2	255, 336, 433, 467, 554, 763, 835, 893, 953	14.0	22.75	134.54	6.23	88.6	22.4	6852
970310	11:28	WFPC2	255, 336, 433, 467, 554, 763, 835, 893, 953	14.0	22.75	204.07	6.08	88.7	22.4	6852
970310	17:55	WFPC2	255, 336, 433, 467, 554, 763, 835, 893, 953	14.0	22.76	298.48	5.87	88.8	22.4	6852
030308	13:54	ACS	226, 271, 336, 343, 431, 477, 502, 658, 891	6.4	-4.09	9.13	38.53	148.7	48.7	9384
030821	11:15	WFPC2	255, 334, 409, 501, 589, 628, 673, 750, 855, 953, 1042	25.0	-18.96	195.99	8.13	245.3	12.5	9738
030822	4:51	WFPC2	255, 334, 409, 501, 589, 628, 673, 750, 855, 953, 1042	25.0	-18.94	93.50	7.63	245.8	12.5	9738
030826	22:32	WFPC2	255, 334, 409, 501, 589, 628, 673, 750, 855, 953, 1042	25.1	-18.81	316.91	5.13	248.8	12.4	10065
030827	9:45	WFPC2	255, 334, 409, 501, 589, 673, 750, 855, 953, 1042	25.1	-18.81	121.03	5.00	249.1	12.4	10065
030828	3:23	WFPC2	255, 334, 409, 501, 589, 628, 673, 750, 855, 953, 1042	25.1	-18.79	19.05	4.88	249.5	12.4	9738

[1]Read 970310 as March 10, 1997

[2]Time given as the start of the ∼25 to 50 minute observing sequence

[3]Resolution is the best spatial resolution at the sub-Earth point for images obtained on the PC1 chip for WFPC2 or the HRC chip for ACS

[4]Space Telescope Science Institute Program Identification number, for HST data archive access

Table 4.4: Adopted Broadband Radiance Estimation Coefficients, ω

Band Code[1]	Band Combination	Coefficients, ω_k				
		$k = 1$ (425 nm)	$k = 2$ (540 nm)	$k = 3$ (654 nm)	$k = 4$ (749 nm)	$k = 5$ (860 nm)
1	5					0.511
2	1	0.424				
3	1, 5	0.045				0.460
4	3			0.134		
5	3, 5			-0.003		0.524
6	1, 3	0.090		0.107		
7	1, 3, 5	0.073		0.070		0.157
8	4				0.364	
9	4, 5				-0.015	0.532
10	1, 4	0.160			0.235	
11	1, 4, 5	0.056			0.035	0.398
12	3, 4			0.138	-0.011	
13	3, 4, 5			0.002	-0.016	0.526
14	1, 3, 4	0.096		0.089	0.043	
15	1, 3, 4, 5	0.086		0.071	0.036	0.092
16	2		0.154			
17	2, 5		0.047			0.355
18	1, 2	-0.037	0.167			
19	1, 2, 5	0.010	0.042			0.361
20	2, 3		0.067	0.076		
21	2, 3, 5		0.049	0.016		0.288
22	1, 2, 3	0.058	0.031	0.090		
23	1, 2, 3, 5	0.037	0.033	0.047		0.182
24	2, 4		0.102		0.127	
25	2, 4, 5		0.059		0.045	0.255
26	1, 2, 4	0.033	0.086		0.137	
27	1, 2, 4, 5	0.024	0.049		0.056	0.244
28	2, 3, 4		0.076	0.045	0.062	
29	2, 3, 4, 5		0.059	0.006	0.043	0.236
30	1, 2, 3, 4	0.057	0.041	0.060	0.060	
31	1, 2, 3, 4, 5	0.046	0.041	0.042	0.053	0.090

[1]For an exposure in which all of the bands in an image are being read out, the filter path code (described in the text) of the band furthest from readout is equal to the "band code."

1. 8-bit to 11-bit decoding

The 8-bit-per-pixel square-root encoded EDR data is decoded to its full 11-bit-per-pixel linear range using the inverse look-up table shown in Table 4.5. The decoded values can range from 0 to 2040.

2. "Bad" pixel identification

Bad pixels, also known as null pixels, are assigned the value specified by the CORE_NULL keyword in the PDS RDR label, and are ignored in all subsequent THEMIS-VIS processing. Bad pixels are identified in the decoded (11-bit-per-pixel) EDR data as follows:

a. Threshold values All pixels with a DN level of 2040 (the highest possible value) or 0 are flagged as null pixels because they are likely saturated.

b. Bad rows and columns Some rows and columns near the edge of each framelet are always filled with unusable (noisy or saturated or zero) data. The pixels in these rows and columns (Table 4.6) are therefore flagged as null.

c. Saturated regions In exposures where all or part of the detector is saturated, the DN value of some saturated pixels is "wrapped" by the instrument firmware to a value which is invariably small compared to the rest of the pixels in the array. These "wrapped" pixels are identified by considering each THEMIS-VIS framelet separately. Any pixel whose value is less than the framelet median by at least 1200 DN has probably been "wrapped" and is therefore flagged as null. The median is calculated using all pixels not flagged as null in **a.** or **b.**.

d. Neighboring pixels Pixels near to saturated regions are often observed to be anomalously high even if they themselves do not meet the criteria of **a.** or **c.**. Therefore, if too many of the pixels near a given pixel are flagged as null in **a.** or **c.**, that pixel is also flagged as null. A 5 by 5 pixel region centered on each pixel is considered. This region is truncated if it would extend pass the edge of the array. Pixels that lie in the "bad rows and columns" identified in **b.** are always treated as valid for this procedure. If more than 30% of the pixels in the 5x5 (truncated if necessary) region for a given pixel are null, then that pixel is also flagged as null.

3. Bias subtraction

a. Nighttime VIS images The bias subtraction procedure uses THEMIS-VIS images acquired at night with nominal exposure times of zero. The bias is determined independently for each of the three available spatial summing modes. In practice there is no detectable difference between night images acquired with typical VIS exposure times (up to thirty milliseconds) and night images with zero exposure time, so non-zero exposure time night images are also included in the averaging process that is used to determine the bias. This also means that the dark current for typical THEMIS-VIS exposure times is *effectively* zero. (Typical on-orbit focal plane temperatures are $\pm 5^{\circ}$C). However, because the readout time is much longer than the exposure times, some of the signal for zero millisecond nighttime exposures may be attributable to dark current accumulated during readout. The evidence for this is a slight ramp-up in the zero-millisecond signal towards the top of the detector. Fortunately, since the amount of dark current that accumulates during readout is not a function of exposure time, and since we have detected no temporal variability in the zero exposure night-time images, there is

Table 4.5: 8-bit to 11-bit Lookup Table

8 bit	11 bit	**8 bit**	11 bit	**8 bit**	11 bit	**8 bit**	11 bit	**8 bit**	11 bit	**8 bit**	11 bit	**8 bit**	11 bit	**8 bit**	11 bit
0	0	**32**	45	**64**	150	**96**	315	**128**	542	**160**	829	**192**	1177	**224**	1586
1	1	**33**	47	**65**	154	**97**	321	**129**	550	**161**	839	**193**	1189	**225**	1599
2	2	**34**	50	**66**	158	**98**	328	**130**	558	**162**	849	**194**	1201	**226**	1613
3	3	**35**	52	**67**	163	**99**	334	**131**	566	**163**	859	**195**	1212	**227**	1627
4	3	**36**	55	**68**	167	**100**	340	**132**	574	**164**	869	**196**	1225	**228**	1641
5	4	**37**	57	**69**	171	**101**	346	**133**	582	**165**	879	**197**	1237	**229**	1655
6	5	**38**	60	**70**	176	**102**	353	**134**	591	**166**	889	**198**	1249	**230**	1669
7	5	**39**	63	**71**	181	**103**	359	**135**	599	**167**	900	**199**	1261	**231**	1683
8	6	**40**	65	**72**	185	**104**	366	**136**	608	**168**	910	**200**	1273	**232**	1697
9	7	**41**	68	**73**	190	**105**	373	**137**	616	**169**	920	**201**	1286	**233**	1712
10	8	**42**	71	**74**	195	**106**	379	**138**	625	**170**	931	**202**	1298	**234**	1726
11	9	**43**	74	**75**	200	**107**	386	**139**	633	**171**	941	**203**	1310	**235**	1740
12	10	**44**	77	**76**	205	**108**	393	**140**	642	**172**	952	**204**	1323	**236**	1755
13	11	**45**	80	**77**	210	**109**	400	**141**	651	**173**	963	**205**	1336	**237**	1769
14	13	**46**	83	**78**	215	**110**	407	**142**	660	**174**	973	**206**	1348	**238**	1784
15	14	**47**	86	**79**	220	**111**	414	**143**	669	**175**	984	**207**	1361	**239**	1798
16	15	**48**	90	**80**	225	**112**	421	**144**	678	**176**	995	**208**	1374	**240**	1813
17	17	**49**	93	**81**	230	**113**	428	**145**	687	**177**	1006	**209**	1386	**241**	1828
18	18	**50**	96	**82**	235	**114**	435	**146**	696	**178**	1017	**210**	1399	**242**	1842
19	20	**51**	100	**83**	241	**115**	442	**147**	705	**179**	1028	**211**	1412	**243**	1857
20	21	**52**	103	**84**	246	**116**	449	**148**	714	**180**	1039	**212**	1425	**244**	1872
21	23	**53**	107	**85**	251	**117**	457	**149**	723	**181**	1050	**213**	1438	**245**	1887
22	25	**54**	110	**86**	257	**118**	464	**150**	732	**182**	1061	**214**	1451	**246**	1902
23	26	**55**	114	**87**	262	**119**	472	**151**	742	**183**	1073	**215**	1464	**247**	1917
24	28	**56**	118	**88**	268	**120**	479	**152**	751	**184**	1084	**216**	1478	**248**	1932
25	30	**57**	121	**89**	274	**121**	487	**153**	761	**185**	1095	**217**	1491	**249**	1947
26	32	**58**	125	**90**	279	**122**	494	**154**	770	**186**	1107	**218**	1504	**250**	1963
27	34	**59**	129	**91**	285	**123**	502	**155**	780	**187**	1118	**219**	1518	**251**	1978
28	36	**60**	133	**92**	291	**124**	510	**156**	789	**188**	1130	**220**	1531	**252**	1993
29	38	**61**	137	**93**	297	**125**	518	**157**	799	**189**	1142	**221**	1545	**253**	2009
30	40	**62**	141	**94**	303	**126**	526	**158**	809	**190**	1153	**222**	1558	**254**	2024
31	43	**63**	145	**95**	309	**127**	534	**159**	819	**191**	1165	**223**	1572	**255**	2040

Table 4.6: Bad rows and columns

Summing Mode	Bad Columns	Bad Rows
1	0 – 9, 1000 – 1023	0 - 1
2	0 – 4, 500 – 511	0
4	0 – 1, 250 – 255	0

Table 4.7: THEMIS-VIS images used for Fig. 4.17 comparison with concurrent HST images

Image Parameters					Image Center Coords.		
Image number	Start time (UT)	Exposure duration (msec)	Summing Mode	Number of framelets	East Lon. (°)	Lat. (°)	Incidence Angle (°)
V05454015	2003-03-08 13:39:00.640	2.8	4	63	43.11	71.34	76.73
V05454017	2003-03-08 13:44:27.441	5.2	2	15	37.26	56.35	74.80
V05454018	2003-03-08 13:46:29.941	5.2	2	15	35.82	50.25	74.31
V05454019	2003-03-08 13:48:19.038	2.7	4	63	34.49	43.62	73.99
V05454021	2003-03-08 13:53:08.639	3.5	2	15	32.26	30.34	74.01
V05454022	2003-03-08 13:56:54.338	7.3	1	3	30.63	19.32	74.70
V05454023	2003-03-08 13:58:25.639	3.0	2	15	29.94	14.43	75.20
V05454024	2003-03-08 14:00:17.838	3.7	2	15	29.17	8.79	75.90
V05454026	2003-03-08 14:02:14.837	3.2	2	15	28.37	2.89	76.78
V05455001	2003-03-08 14:04:29.439	2.2	4	63	27.28	-5.12	78.20
V05455003	2003-03-08 14:06:52.837	4.8	2	15	26.46	-11.16	79.42
V05455005	2003-03-08 14:07:52.138	5.2	2	15	26.05	-14.16	80.07
V05455006	2003-03-08 14:09:45.036	6.2	2	15	25.25	-19.89	81.39
V05455007	2003-03-08 14:12:56.836	8.8	2	15	23.82	-29.64	83.80
V05455008	2003-03-08 14:15:27.137	7.9	4	63	22.40	-38.51	86.15
V05455023	2003-03-08 15:37:43.430	3.1	4	63	13.92	70.92	76.65
V05455025	2003-03-08 15:41:35.726	3.0	4	63	9.22	59.44	75.10
V05455026	2003-03-08 15:44:15.327	5.2	2	15	7.48	52.70	74.48
V05455027	2003-03-08 15:46:33.429	4.7	2	15	6.04	45.83	74.07
V05455028	2003-03-08 15:48:44.725	4.8	2	15	4.84	39.28	73.90
V05455029	2003-03-08 15:50:36.526	5.2	2	15	3.92	33.69	73.92
V05455030	2003-03-08 15:52:35.127	4.8	2	15	3.00	27.76	74.12
V05455031	2003-03-08 15:54:38.428	8.2	1	3	2.14	21.88	74.49
V05455032	2003-03-08 15:56:40.928	2.0	4	63	1.06	14.23	75.22
V05455033	2003-03-08 15:59:41.025	3.5	2	15	359.99	6.37	76.25
V05455034	2003-03-08 16:01:26.724	3.8	2	15	359.27	1.04	77.10
V05456001	2003-03-08 16:03:35.927	4.4	2	15	358.39	-5.48	78.28
V05456002	2003-03-08 16:06:09.126	5.3	2	15	357.34	-13.23	79.88
V05456003	2003-03-08 16:09:29.626	7.4	2	15	355.91	-23.40	82.25
V05456004	2003-03-08 16:13:12.125	6.9	4	63	354.01	-35.94	85.47
V07470002	2003-08-21 10:55:57.821	1.6	2	15	236.59	-7.51	74.40
V07470005	2003-08-21 11:00:51.618	3.2	1	3	234.55	-22.12	68.75
V07470007	2003-08-21 11:02:48.418	1.5	2	15	233.62	-28.37	66.71
V07470008	2003-08-21 11:03:50.320	1.5	2	15	233.14	-31.52	65.77
V07470009	2003-08-21 11:04:59.719	2.0	2	15	232.57	-35.06	64.81
V07470010	2003-08-21 11:05:57.117	0.9	4	63	231.88	-39.22	63.80
V07470011	2003-08-21 11:07:50.019	1.9	2	15	231.07	-43.76	62.86
V07470012	2003-08-21 11:08:49.718	2.0	2	15	230.48	-46.81	62.33
V07470013	2003-08-21 11:09:51.320	2.1	2	15	229.83	-49.96	61.87
V07470014	2003-08-21 11:11:04.417	4.2	1	3	229.06	-53.39	61.47
V07470015	2003-08-21 11:11:45.320	2.1	2	15	228.47	-55.79	61.26
V07470016	2003-08-21 11:12:46.019	2.1	2	15	227.63	-58.89	61.07
V07470017	2003-08-21 11:13:49.718	2.2	2	15	226.62	-62.15	60.96
V07470018	2003-08-21 11:14:53.417	1.9	2	15	225.42	-65.40	60.97
V07470019	2003-08-21 11:15:58.417	1.4	4	63	223.28	-69.94	61.14
V07479007	2003-08-22 04:48:09.321	3.1	2	15	334.78	-23.70	68.13
V07479008	2003-08-22 04:49:09.723	5.5	1	3	334.38	-26.47	67.22
V07479009	2003-08-22 04:49:46.622	2.9	2	15	334.05	-28.65	66.54
V07479010	2003-08-22 04:50:47.223	2.8	2	15	333.58	-31.73	65.63
V07479011	2003-08-22 04:51:46.821	2.6	2	15	333.10	-34.77	64.80
V07479012	2003-08-22 04:52:56.223	2.3	2	15	332.52	-38.31	63.93
V07479013	2003-08-22 04:53:57.223	2.5	2	15	331.99	-41.43	63.24
V07479014	2003-08-22 04:54:57.820	2.5	2	15	331.43	-44.52	62.65
V07479015	2003-08-22 04:55:59.020	1.4	4	63	330.58	-48.87	61.94
V07479016	2003-08-22 04:57:49.019	2.5	2	15	329.62	-53.27	61.41
V07479017	2003-08-22 04:58:49.422	2.5	2	15	328.86	-56.36	61.15
V07479018	2003-08-22 05:00:34.722	2.5	2	15	327.32	-61.74	60.91
V07479019	2003-08-22 05:01:37.222	1.4	4	63	325.71	-66.15	60.93
V07479020	2003-08-22 05:03:24.120	1.2	4	63	322.93	-71.59	61.21
V07479021	2003-08-22 05:05:10.320	1.0	4	63	318.34	-76.95	61.78
V07479022	2003-08-22 05:06:54.120	1.5	2	15	311.77	-80.89	62.39
V07479023	2003-08-22 05:07:53.421	1.0	4	63	295.25	-84.71	63.19
V07536016	2003-08-26 23:11:13.388	3.8	4	63	106.98	11.58	82.56
V07536017	2003-08-26 23:13:04.790	4.4	2	15	106.37	7.18	80.44
V07536018	2003-08-26 23:13:59.188	3.9	2	15	106.00	4.44	79.14
V07542009	2003-08-27 09:36:06.631	6.0	4	63	267.19	-85.73	62.85
V07551003	2003-08-28 03:10:12.835	0.9	4	63	54.27	-47.97	61.04
V07551007	2003-08-28 03:16:36.635	1.5	2	15	49.02	-66.35	59.95

no need to distinguish between bias and readout dark current for the purpose of calibration. We will therefore henceforth lump both effects together and refer to the combination of the two simply as "bias."

b. Averaging to create bias frames To properly account for the small ramp in the bias level, bias frames are created for each of the 31 filter paths and for each of the 3 spatial summing modes. Where possible, each bias frame is generated from a simple average of all available nighttime exposure framelets with the same summing mode and filter path.

c. Modeling to handle paths for which no empirical data is available For many of the less common filter paths, nighttime exposure framelets have not yet been acquired. In these cases, we apply the following simple model of the detector readout process in order to approximate the expected bias signal:

Consider what happens to the charge from all upstream filters while a downstream filter is being read out. As our first simplifying assumption, we ignore the inter-framelet rows (there are 7 ± 2 inter-framelet rows at each filter boundary) and pretend that the top row of one framelet is always adjacent to the bottom row of the next. As each row of the downstream filter is shifted into the h-register, the upstream filter rows are shifted one row closer to the h-register, so that by the time the entire downstream filter has been read out, the rows of each upstream filter f have shifted so that they now lie under the filter $f - 1$. The charge that these rows accumulate in that time is characteristic of filter f. When the next downstream filter is read out, these same rows are shifted from lying under filter $f - 1$ to lying under filter $f - 2$, with the charge accumulated during the shift being characteristic of the starting filter for that shift. When the rows of filter f

are finally in the position where filter f rows are being transferred directly to the h-register, the charge that they accumulate is characteristic of filter 1. Thus, if all of the filters downstream from f are read out in an exposure, then the rows in filter f accumulate charge from filters f, $f-1$, $f-2$, $\ldots 1$.

If any of the filters downstream from f is not being read out, then the process of dumping that filter shifts f downstream very rapidly, so if the rows in filter f were under filter $f-\Delta$ at the start of the dump, very little of the characteristic bias charge of filter $f-\Delta$ is accumulated. Thus, our second simplifying assumption is that the charge accumulated during such a dump is negligible.

The bias charge built up during readout for a certain filter following a given filter path with filter path code F can therefore be calculated simply by summing the characteristic bias charges, $E_{ij,f}$, of each downstream filter. The bias frame $B_{ij,F}$ is:

$$B_{ij,F} = \sum_{f=1}^{f_0} \eta(f) \cdot E_{ij,f_0-f+1} \; . \tag{4.11}$$

The values of f_0 and $\eta(f)$ are properties of the filter path denoted by F and are defined by Eq. (4.4), *i.e.*, f_0 is the starting filter of the filter path and $\eta(f)$ indicates whether the filter f is one of those being read out for that filter path. $E_{ij,f}$ and $B_{ij,F}$ are elements of arrays with the same dimensions as the framelets of the spatial summing mode for which the $B_{ij,F}$ are being derived. Note that $\eta(f)$ is always equal to 1 when $f = f_0$, so the characteristic bias of the first filter, $E_{ij,f=1}$, is always present in the bias frame.

$E_{ij,f}$ can be derived by using empirically known (via the averaging method described above in item **b.**) Bij, F. We perform this derivation by using $B_{ij,F}$ values for those filter paths for which all downstream filters are read out. These are $B_{ij,F=31}$, $B_{ij,F=15}$, $B_{ij,F=7}$, $B_{ij,F=3}$, and $B_{ij,F=1}$. Writing out Eq. (4.11) explicitly

for these five Bij,F leads to the following trivial system of equations:

$$\begin{aligned} B_{ij,F=31} &= E_{ij,f=1} + E_{ij,f=2} + E_{ij,f=3} + E_{ij,f=4} + E_{ij,f=5} \\ B_{ij,F=15} &= E_{ij,f=1} + E_{ij,f=2} + E_{ij,f=3} + E_{ij,f=4} \\ B_{ij,F=7} &= E_{ij,f=1} + E_{ij,f=2} + E_{ij,f=3} \\ B_{ij,F=3} &= E_{ij,f=1} + E_{ij,f=2} \\ B_{ij,F=1} &= E_{ij,f=1} \end{aligned} \tag{4.12}$$

d. Bias subtraction algorithm Bias frames for each summing mode are stored in FITS files each with 31 planes corresponding to the 31 filter paths. The bias is removed by determining the filter path for each framelet of each filter and then subtracting from it the stored bias frame for that filter path and summing mode. The bias frames are stored in the FITS files in order of the F code, *i.e.*, the bias frame for a given filter path is found at plane number $F-1$ for a zero-based cube plane counting system.

4. Register stray light subtraction

The strategy for removing register stray light is to first identify the spatial pattern using zero duration exposures, and then determine the way that the intensity of that pattern scales with C-ROI mean broadband scene radiance $\widehat{\overline{I}}$.

a. Zero duration daytime THEMIS-VIS images On-orbit images acquired during daylight with a commanded exposure time of zero msec show the same spatial pattern as the ground-calibration-derived zero exposure images. (*cf.*, Fig. 4.9 and Fig. 4.15) Of course, a commanded exposure time of zero does not really produce an exposure time of exactly zero. However, in the majority of zero exposure

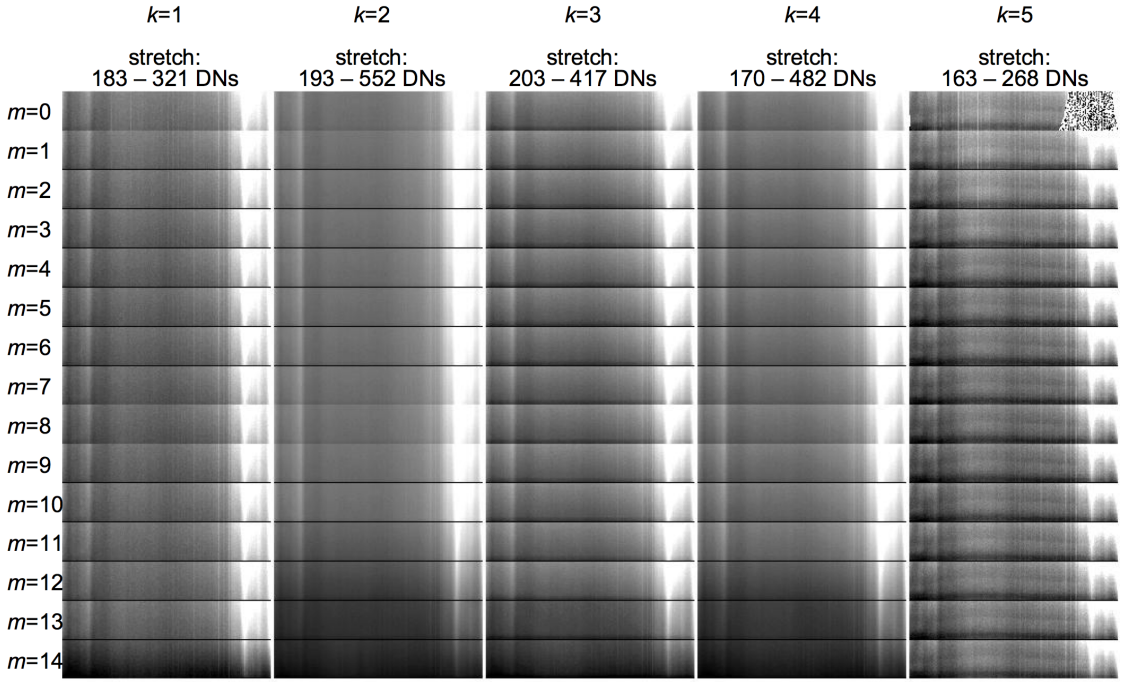

Figure 4.15: A THEMIS-VIS EDR, as in Figure 4.4, but of a zero-exposure-time in-flight image.

flight data, there is no evidence of surface features visible through the register stray light pattern. The spatial variation of the small amount of scene derived DN that is probably present in the zero exposure images will be washed out by averaging a sufficient number of images. Furthermore, the level of scene-derived contamination of the zero exposure images will be a negligible fraction of the signal of any given THEMIS-VIS image as long as the effective exposure time for a commanded exposure time of zero is a negligible fraction of the effective exposure time for the THEMIS-VIS image in question.

The register stray light frames are generated using zero duration daytime THEMIS-VIS images that have been calibrated through stage 3 (bias subtraction)

of the pipeline. The zero exposure daytime images are first averaged according to filter path and spatial summing mode in the same manner as the nighttime images are averaged to create bias frames. Next, the register stray light frames are normalized by, first, setting the mean of a given spatial summing mode's $F = 1$ register stray light frame equal to unity. Then, the register stray light frame for every other filter path for that spatial summing mode is normalized so that its values, relative to the $F = 1$ frame, reflect the relative intensity of the register stray light for that filter path in the data set. Normalization to reflect relative intensities in the data set is crucial, because the only means that we have available for determining the scaling factor between the broadband radiance $\widehat{\overline{I}}$ and register stray light is to measure the changes in mean DN from framelet to framelet within an image sequence as a result of changes in filter path, and then find a single scaling factor for all of the register stray light frames that eliminates these changes in mean DN.

For each filter path, the necessary normalization factor is determined as follows: 1) Find all of the exposures that are in the set of images used to generate the register stray light frames, and contain both the given filter path and the $F = 1$ filter path. 2) For each of these exposures, calculate the ratio of the mean value of the framelet from the given filter path to the mean value of the $F = 1$ framelet. 3) The normalization factor is a single number by which the filter path's register stray light frame is scaled so that its mean equals the mean of the ratios from the previous step. Note that the means used to normalize the register stray light frames are calculated from the all non-null pixels in the framelets, rather than from the C-ROI used elsewhere in the calibration procedures.

A complication for this ratio-based normalization is that there are many filter

paths that cannot occur in the same exposure as the $F = 1$ path. However, every filter path *does* occur in the same exposure as one of $F = 16$, $F = 8$, $F = 4$, $F = 2$, or $F = 1$ paths, which we will refer to as "clear" paths (these are the paths that apply to filters 5, 4, 3, 2, and 1, respectively, when there are no other filters downstream). The normalization factors of the "clear" filter paths are established using ratios with respect to the $F = 1$ path, according to the previously described procedure. Then the ratios of all other filter paths are measured with respect to whichever "clear" path they occur with, and then rescaled using the ratio of that "clear" path to the $F = 1$ path, which gives us the ratios of these paths with respect to the $F = 1$ path and hence the normalization factor.

The "clear"-to-$F = 1$ ratios are established as follows: In THEMIS-VIS images that use all five filters, an exposure containing $F = 2$ always occurs immediately following an $F = 1$ exposure, an exposure containing $F = 4$ always occurs immediately following an $F = 2$ exposure, etc. So the $F = 2$ ratio is calculated relative to the $F = 1$ from the previous exposure in the same image, the $F = 4$ is calculated relative to the previous exposure $F = 2$, $F = 8$ relative to the previous $F = 4$, and $F = 16$ relative to the previous $F = 8$. These ratios are then rescaled in sequence so that they are expressed relative to $F = 1$.

This normalization scheme assumes that the register stray light signals from the various filter paths are related to each other by a constant proportionality factor. We have examined the register stray light signal ratios, and have found that they are consistent with the constant proportionality hypothesis. Obviously, the normalized register stray light frames are generated as described above only for the filter paths that are represented in our data set of zero exposure daylight images. Register stray light frames for filter paths that are not represented in

the data set of zero exposure daylight images are generated by starting from the normalized frames and then applying the same method used for the bias frames in step **3c**.

b. Deriving the register stray light response coefficient In order to solve for the register stray light response coefficient, we postulate that, *on average*, once the register stray light is removed, the framelet means will be independent of the filter path, provided that we consider framelets from the same filter. To take advantage of this postulate, we first determine, for each 5-filter image of a given summing mode, the factor ψ that we need to multiply the normalized register stray light frames by so that when they are subtracted from the image framelets, the differences between adjacent framelet C-ROI means are minimized in a least-squares sense. More explicitly, for each 5-filter image, we: 1) Identify the set of framelet pairs that satisfy: a) both members of the pair are from the same filter; b) the members of the pair come from adjacent (in the time sequence) exposures; and c) the members of the pair have different filter paths. Let the elements of this set be indexed by u. 2) Calculate the difference in the mean (bias subtracted) DN for each pair, $\overline{d}_u$, and the difference in the mean normalized register stray light frame for the filters paths of the members of each pair, $\overline{g}_u$. 3) Find the value of ψ that minimizes

$$\sum_u \left(\psi \overline{g}_u - \overline{d}_u\right)^2 . \tag{4.13}$$

4) Record ψ, and each band's mean photosite signal $\overline{S}_k$, for the image. If $\overline{G}_{u'}$ is the C-ROI-mean of the normalized register stray light frame for a framelet pair member's filter path, and $\overline{D}_{u'}$ is the C-ROI-mean uncorrected DN for that pair

member, and t is the exposure time, then

$$t\overline{S}_k = \underset{k'=k}{\text{mean}} \left(\overline{D}_{u'} - \psi \overline{G}_{u'} \right) \ . \tag{4.14}$$

The mean is over all elements of the u' set of filter pair members for which the filter number is equal to the k in the $\overline{S}_k$ we are calculating.

The register stray light response coefficient z is the factor that relates the mean broadband scene radiance $\widehat{\overline{I}}$ to ψ:

$$\psi = z \, \widehat{\overline{I}} \ . \tag{4.15}$$

Solving for ψ by minimizing Eq. (4.13) would of course be sufficient to remove register stray light in a THEMIS-VIS sequence of a uniform scene, but in order to be able to calibrate any arbitrary sequence, we need to know z. To estimate z we select a sample of 5-band images with low scene variance and determine ψ and the $\overline{S}_k$ as described above. Let the elements of this sample be indexed by v, so we have a set of ψ_v and $\overline{S}_{k,v}$.

The $\overline{S}_{k,v}$ yield $\widehat{\overline{I}}_v$ by applying Eq. (4.9), and we use the resulting set of ψ_v and $\widehat{\overline{I}}_v$ with Eq. (4.15) to give a least squares solution for z. In finding the $\widehat{\overline{I}}_v$, we are free to chose which of the bands k of the $\overline{S}_{k,v}$, and thus which set of, ω_k to use. Naively, we might expect that using all five bands would give the best estimates and thus the smallest residuals on the z solution, because it provides the most information about the scene radiance spectrum. However, $\psi\overline{G}$ is an imperfect solution for the register stray light, and so in general the $\overline{S}_k$ estimates will be contaminated by some register stray light, reducing the accuracy of the $\widehat{\overline{I}}_v$. This contamination is least significant in the bands with the highest photosite signal, and thus we find that using only the $k = 3$ (650 nm) band gives the fit for z with the highest R^2. The $k = 3$ solution is our adopted z value, and is reported in Table 4.2. Excluding

the $k = 3$, 4 band combination, which gives solutions and R^2 identical to that of $k = 3$, the band combination with the next highest R^2 is $k = 2, 3$. Since the difference in the z solutions of these two best fits is much higher than their formal errors, we use this difference to form our confidence interval, with the half-width of the confidence interval being set equal to the magnitude of $z(k = 1) - z(k = 2,\ 3)$.

c. Subtraction of scaled register stray light frames Register stray light frames for each summing mode are stored in FITS files each with 31 planes corresponding to the 31 filter paths. These frames are stored in order of the F code, *i.e.*, the frame for a given filter path is found at plane number $F - 1$ for a zero-based cube plane counting system.

Let D_{ijfa} be the bias-subtracted (*i.e.*, calibrated through stage 3) DN for column i, row j, filter f, and exposure a of a THEMIS-VIS imaging sequence with exposure duration t, and let $G_{ij,F(f,a)}$ be the i, j elements of the normalized register stray light frame with the appropriate filter path F. (F is determined from Eq. (4.4).) Then the photosite signal S_{ijfa} is

$$S_{ijfa} = \frac{D_{ijfa} - z\ \widehat{\overline{I}}_a\ G_{ij,F(f,a)}}{t}\ . \tag{4.16}$$

The means of determining broadband C-ROI mean radiance $\widehat{\overline{I}}_a$ is apparent if we take the C-ROI means of S, D, and G, in Eq. (4.16) and then substitute into Eq. (4.9), eliminating S:

$$\widehat{\overline{I}}_a = \frac{\sum_f \omega_f \dfrac{\overline{D}_{f,a'(a,f)}}{t}}{1 + \sum_f \omega_f \dfrac{z\overline{G}_{F(f,a'(a,f))}}{t}}\ . \tag{4.17}$$

The sums are over whatever subset of filters we chose to use, and the ω_f are those appropriate to that subset. The exposure number a' used to calculate the

broadband radiance estimate for exposure number a is *not* in general equal to a, because the source region for an exposure's register stray light may not be within its field of view. Since 50% or more of the register stray light is contributed by the h-register, and since the h-register is at the bottom of the array where it is most likely contaminated by scene radiance near the bottom of an exposure's field of view, we chose a' to select a filter from a different exposure (a later exposure in the sequence) that covers the region just below the field of view of exposure a. We select that later exposure as follows:

$$a'(a,f) = \begin{cases} a & \text{if } f = 1 \\ a+1 & \text{if } f = 2 \\ a+f & \text{if } g \geq 3 \end{cases} . \tag{4.18}$$

Although in principle we might wish to use as many of the THEMIS-VIS filters as possible when calculating $\widehat{\overline{I}}_a$ from Eq. (4.17), in practice we have found that using choosing a single filter produces better results, in the sense that when the results of single-filter-$\widehat{\overline{I}}_a$ register stray light removal and multi-filter-$\widehat{\overline{I}}_a$ register stray light removal are visually inspected and compared, the multi-filter-$\widehat{\overline{I}}_a$ image is more likely to be the one with more prominent register stray light artifacts. Therefore, we calculate $\widehat{\overline{I}}_a$ using D and G from filter 3 (band 3) whenever it is present in the image. If filter 3 is not present, we use filter 4, and if not 4, then the next most preferable filter, with the full preference order being 3, 4, 5, 2, 1. Less preferable filters are those that have higher residuals when we fit for z in Eq. (4.15) using single-filter-derived values for $\widehat{\overline{I}}$.

The series of $\widehat{\overline{I}}_a$ derived using Eq. (4.17) for an image will have missing elements on the edges due to the beginning and end of the exposure sequence, and sometimes in the middle due to invalid image data. Interior gaps are always filled by linear

interpolation. Extrapolation is allowed for only one element past the edge of the valid $\widehat{\overline{I}}_a$ elements, and $\widehat{\overline{I}}_a$ is treated as constant from this edge point and outwards.

5. Correction for pixel response variation ("flatfielding")

The response of framelet pixels to scene radiance is variable, due both to variations in the sensitivity of detector pixels, and to variations in optical throughput. In the absence of stray light, an image of a uniform scene is a "flatfield," *i.e.*, a map of these response variations.

Unfortunately, the variability in THEMIS-VIS signal in uniform-scene images is dominated by stray light. The crucial distinction between response variations and stray light is that response variations are a multiplicative effect, while stray light is additive, so that regions of the detector that are brighter due to greater response will have proportionally enhanced contrast, while regions that are brighter due to more stray light will have not have enhanced contrast. This means that if we incorrectly attribute all of the non-uniform signal in a uniform scene image (after removal of register stray light) to photosite stray light, then regions of high response will have spuriously high contrast. We must therefore remove response variations prior to deriving and removing photosite stray light light. However, if we incorrectly attribute all of the non-uniform signal in a uniform-scene image to response variations, our calibration process will have the converse effect of creating spuriously decreased contrast in regions of high stray light. We have therefore adopted a method for generating flatfields that relies on the scene contrast directly by using framelet to framelet signal *differences* rather than averages as in the usual flatfielding approach.

A flatfield is derived independently for each of the five filters using data cali-

brated through stage 4 (register stray light subtraction). Let $S'_{ij,n}$ and $S''_{ij,n}$ be the stage 4 signal for, respectively, the early and later members of framelet pair n. Let $\sigma'_{ij,n}$ and $\sigma''_{ij,n}$ be the standard errors for the DN values of the members of the pair. These standard errors are based on Poisson counting noise statistics, *i.e.*, they are simply the square root of the numbers of electrons detected at each pixel, so:

$$\sigma = \frac{\sqrt{(\text{electrons per DN}) \cdot (\text{raw DN level})}}{(\text{electrons per DN})} \,. \tag{4.19}$$

The "raw DN level" refers to the DN level measured prior to bias subtraction, *i.e.*, calibrated through stage 2. The gain is 25.4 electrons per DN for spatial summing mode 1, 25.4/2 for spatial summing mode 2, and 25.4/4 for spatial summing mode 4. A framelet pair consists of any two adjacent framelets from the same filter and same THEMIS-VIS data cube. Let the total number of such framelet pairs in the data set for a given filter be labeled N. We assume that the photosite stray light signal is slowly varying from framelet to framelet, so that the difference between framelets depends only on the changes in the scene and on counting noise. Since the magnitude of the differences caused by scene changes (as a fraction of the mean signal) is proportional to responsivity, we estimate the responsivity R by starting with the variance of the framelet-to-framelet differences:

$$\frac{1}{N}\sum_{n=1}^{N}\left(\frac{S''_{ij,n} - S'_{ij,n}}{\text{mean}\big(S'_{ij,n}\big)}\right)^2 \,, \tag{4.20}$$

and then subtracting the contribution of counting noise to those differences:

$$\frac{1}{N}\sum_{n=1}^{N}\left(\frac{\sigma''_{ij,n}}{\text{mean}\big(S'_{ij,n}\big)}\right)^2 + \left(\frac{\sigma'_{ij,n}}{\text{mean}\big(S'_{ij,n}\big)}\right)^2 \,. \tag{4.21}$$

The responsivity R is the square-root of the counting-noise corrected variances:

$$R_{ij} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\frac{\big(S''_{ij,n} - S'_{ij,n}\big)^2 - \big(\sigma''_{ij,n}\big)^2 - \big(\sigma'_{ij,n}\big)^2}{\text{mean}\big(S'_{ij,n}\big)^2}} \,. \tag{4.22}$$

Unfortunately, the assumption of slowly varying photosite stray light is not always valid, leading to some contamination of our responsivity estimate by stray light effects. This contamination is most severe where the stray light is most severe, on the right and left edges of all filters, and across the entire filter 1 (band 5, 860 nm). Given the strong column dependence of the stray light, the weak variation of derived responsivity with column in less contaminated regions, and the tendency of the variation with column to mimic that of the stray light, we are not confident that any of the column variation in derived responsivity is real. However, the row dependence of responsivity is much stronger, and not correlated with the stray light patterns. We therefore treat the responsivity as being a function only of detector row, allowing us to derive it by averaging across the central, least contaminated columns. Figure 4.16 (dashed lines) shows a family of profiles along columns of R_{ij}, to illustrate the dominance of variation with row. The strong drop in responsivity below row 15 is due to vignetting by the edge of the filter.

We derived the column-averaged flatfield using summing mode 2 R_{ij} values, because the summing mode 2 data set is much larger than that in the other modes, and so produces a higher quality flatfield. The resulting profile is normalized so that it has unit mean within the C-ROI bounds. This normalization ensures that flatfielding does not on average alter $\overline{S}$. Thus, the flatfield as a function of row, R_j, is:

$$R_j = \frac{\left(\sum_{i \in \mathrm{CROI}} R_{ij} \right)}{\left(\frac{1}{\text{number of elements of } j \in \mathrm{CROI}} \right) \left[\sum_{j \in \mathrm{CROI}} \left(\sum_{i \in \mathrm{CROI}} R_{ij} \right) \right]} . \qquad (4.23)$$

Each column of each framelet is corrected by dividing it by R_j. Since R_j is originally generated using only summing mode 2 data and is thus a 96 element vector,

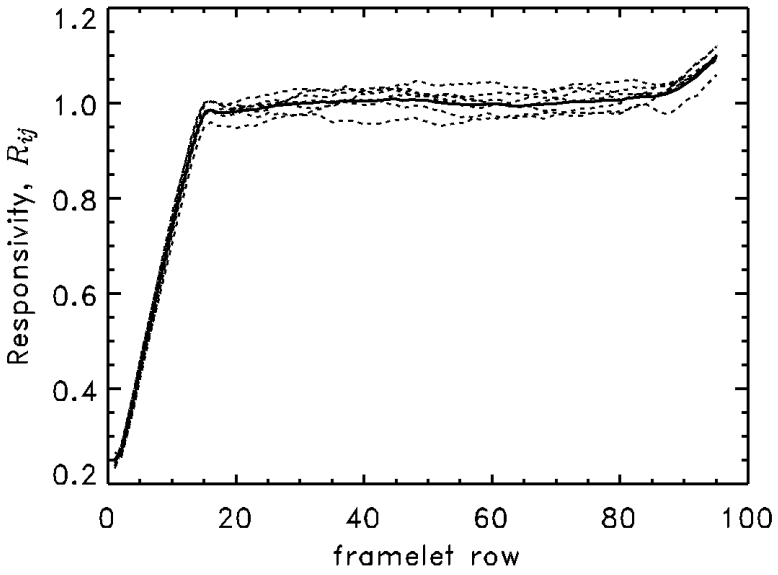

Figure 4.16: Responsivity, R_{ij}, as a function of row for filter $f = 3$ (band $k = 3$), summing mode 2. The dashed lines are generated by averaging groups of 10 columns. These groups of columns are 210–219, 240–249, 270–279, ...390–399. The bold solid line is the adopted solution for R_j. All lines have been scaled by a common normalization factor, such that R_j, has a mean value of 1 over the rows of the C-ROI.

it is resampled to 192 elements via linear interpolation for use with summing mode 1 data, and resampled to 48 elements via pixel averaging for use with summing mode 4 data. The solid line in Fig. 4.16 shows R_j for filter 3, summing mode 2. R_j for filters 2, 4, and 5 is similar. For filter 1, the stray light contamination is so severe that we do not attempt to derive a flatfield, and so R_j for filter 1 is set equal to one for all rows.

6. Photosite stray light subtraction

We model the photosite stray light signal in each framelet as the sum of a spatially variable component $X_{ijk}\,\widehat{\overline{I}}$ and a spatially uniform component $s_k\,\widehat{\overline{I}}$. So $S_{ijk,p}$, the total flatfielded signal (*i.e.*, the signal calibrated through stage 5) for column i, row j, of the p-th band k framelet in the THEMIS-VIS data set, is modeled as:

$$S_{ijk,p} = (X_{ijk} + x_k)\;\widehat{\overline{I}}_p + y_k\left(\overline{I}_{ijk,p} + I'_{ijk,p}\right)\;. \tag{4.24}$$

The parameters x_k and y_k are derived from pre-flight data as previously described and as used in Eq. (4.7). $\overline{I}_{k,p}$ is the mean scene radiance for band k framelet p, also as used in Eq. (4.7), and $I'_{ijk,p}$ represents deviations from that mean scene radiance. All spatial means, denoted by over-bars, are taken over the C-ROI region of each framelet. By combining Eq. (4.24) with Eq. (4.7), we get:

$$S_{ijk,p} - \overline{S}_{k,p} = X_{ijk}\;\widehat{\overline{I}}_p + y_k\,I'_{ijk,p} + \epsilon_{k,p}\;. \tag{4.25}$$

Assuming a sufficiently large sample of framelets, the pattern of radiance in those scenes, $I'_{ijk,p}$, will be essentially random, and so we can combine the $I'_{ijk,p}$ term with the $\epsilon_{k,p}$ error term. Using $S'_{ijk,p}$ to denote deviations from the framelet mean signal, this yields:

$$S'_{ijk,p} = X_{ijk}\;\widehat{\overline{I}}_p + \epsilon_{ijk,p}\;. \tag{4.26}$$

Thus, we solve for X_{ijk} using a least-squares fit of Eq. (4.26). The framelet groups indexed by p are sets of framelets from different bands that come from the same imaging sequence and share the same framelet number m. Each framelet group has a set of $\overline{S}_k$ from which $\widehat{\overline{I}}$ is calculated using Eq. (4.9). X_{ijk} is derived independently for each summing mode using all framelet groups in the THEMIS-VIS data set that contain all five bands. For the fitting process, $\overline{S}_5$, the mean band 5 (860 nm) signal is not used in the calculation of $\widehat{\overline{I}}$.

Using the derived X_{ijk}, determining the corrected photosite DN signal, $Q_{ijk,p}$,

$$Q_{ijk,p} = S_{ijk,p} - (X_{ijk} + x_k)\ \widehat{\overline{I}}_p = y_k I_{ijk,p}\ , \tag{4.27}$$

for any arbitrary framelet group is simply a matter of determining $\widehat{\overline{I}}$ for that group. We do so using Eq. (4.9) with the $\overline{S}_k$ from that group. All valid $\overline{S}_k$ values from a group are used to calculate $\widehat{\overline{I}}$, except that band 5 is excluded, unless band 5 has the only valid value in that framelet group. We exclude band 5 because, by visual inspection of the calibration results, we have found that excluding band 5 leads to a decrease in the prominence of residual photosite stray light artifacts. An $\overline{S}_k$ is valid if at least 50% of the pixels in its C-ROI are non-null. If *all* $\overline{S}_k$ for a framelet group are invalid, then all pixels in that framelet group will be set to null.

7. Conversion to radiance

Given Q_{ijk} from stage 6, and since y_k is derived from pre-flight measurements, I_{ijk}, the calibrated radiance in a THEMIS-VIS framelet, is simply:

$$I_{ijk} = \frac{Q_{ijk}}{y_k}\ . \tag{4.28}$$

4.5 Validation

In order to verify the accuracy of the THEMIS-VIS calibration, we use concurrent observations with HST from 5 different dates in August, 2003, and one date in March, 2003. (We have also made comparisons, which we will described in the next section, with historical HST observations from March, 1997.) For each set of HST observations in 2003 (see Table 4.3 for details), we targeted 5-band THEMIS-VIS images within two hours of the first or last HST image. These THEMIS-VIS images (Table 4.7) were acquired with all three of the available spatial summing modes, and they sample incidence angles of 60° and higher (all THEMIS-VIS emission angles were near zero), and terrains with the full range of martian albedo values.

Figure 4.17 compares radiances derived from the HST image sets with C-ROI mean radiances from the concurrently targeted THEMIS-VIS framelets. This comparison uses only framelets from THEMIS-VIS images with an effective exposure time greater than three milliseconds. (The effective exposure time is the actual exposure time multiplied by the spatial summing mode.) In order to perform the comparison, the HST images are first projected onto a simple cylindrical grid with a 5 pixel-per-degree sampling interval, so that each HST grid box covers 0.2° × 0.2° of latitude and longitude, starting from 90° South, 180° West at the bottom-left corner of the bottom-left grid box. (We use a 3 pixel-per-degree grid for the lower resolution March, 2003 HST images.) We define usable HST grid boxes to be those with both incidence and emission angles of less than 80°. Similarly, we define THEMIS-VIS pixels to be usable if they contain valid data and have incidence and emission angles of less than 80°. If less than 75% of the C-ROI pixels are usable in any band of a framelet, we do not use that framelet.

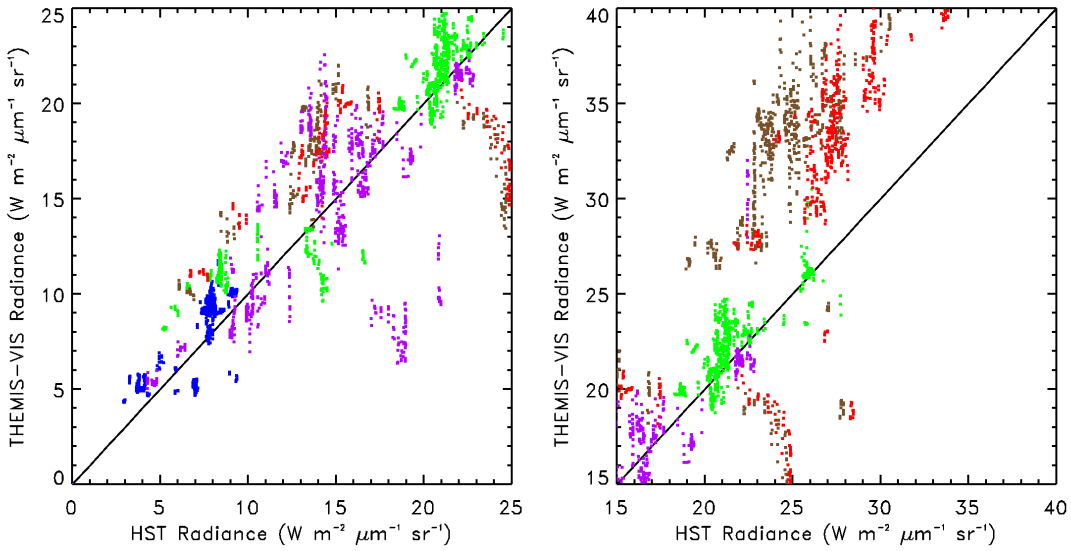

Figure 4.17: The C-ROI mean radiance in THEMIS-VIS framelets, plotted against the photometrically corrected radiance measured in concurrent HST images. The data points are color coded, with blue, green, red, brown, and purple corresponding to bands $k = 1$, 2, 3, 4, and 5 (425, 540, 654, 749, and 860 nm), respectively. A solid line with zero intercept and a slope of one has been drawn in order to highlight the location of points for which the HST and THEMIS-VIS radiances are exactly equal. The left and right-hand plots are identical, except for the range of the axes. None of the $k = 1$ data points fall within the range of the right-hand plot's axes.

To generate HST radiances for the comparison, we do the following separately for each band of each framelet:

1. Find the set of usable C-ROI pixels.

2. Identify the THEMIS-VIS radiance as the mean radiance of the usable C-ROI pixels.

3. Identify the THEMIS-VIS incidence, emission, and phase angles as the mean values of these quantities over the C-ROI pixels.

4. For each narrow-band HST image, identify the HST I/F, incidence angle, and emission angle as the weighted mean of the quantities over the HST grid boxes into which the usable THEMIS-VIS C-ROI pixels fall. The weights are the fraction of usable THEMIS-VIS C-ROI pixels in each grid box. (The HST phase angle is, of course, a constant within each HST image.)

5. Apply a photometric correction to the HST I/F values, using a Minnaert photometric model to adjust them to values appropriate for the THEMIS-VIS observations. The correction factor is:

$$\frac{\beta_{h_{\mathrm{THEMIS}}} \left[\mu_{0,\mathrm{THEMIS}}\right]^{\kappa_{h_{\mathrm{THEMIS}}}} \left[\mu_{\mathrm{THEMIS}}\right]^{\kappa_{h_{\mathrm{THEMIS}}}-1}}{\beta_{h_{\mathrm{HST}}} \left[\mu_{0,\mathrm{HST}}\right]^{\kappa_{h_{\mathrm{HST}}}} \left[\mu_{\mathrm{HST}}\right]^{\kappa_{h_{\mathrm{HST}}}-1}} , \tag{4.29}$$

where μ_0 is the incidence angle, μ is the emission angle, h is the phase angle. The Minnaert coefficients, β and κ (both are functions of the phase angle h) are given in Table 4.8, and were derived using the HST data set and methods described in Soderblom et al. [2005]. The coefficients were derived using narrow-band filters at the wavelengths listed in the table. To correct HST radiances at other wavelengths, we interpolate the coefficients linearly in wavelength, unless the target wavelength is outside of the wavelength range of

tabulated coefficients, in which case we use the coefficients from the nearest-neighbor wavelength. Although separate Minnaert coefficients are used for each phase angle bin, correcting for phase effects is still problematic, because the HST data used to derive the correction do not exceed 40° phase and all of the THEMIS-VIS data used for the radiance comparison have phase angles greater than 60°. No earth-based instruments can observe at phase angles as large as those of THEMIS-VIS in its current nadir-viewing configuration and 5pm orbit.

6. Estimate the spectral radiance by fitting a cubic spline in wavelength space to the HST narrow-band I/F values and then multiplying by the solar spectral radiance.

7. Identify the HST radiance by weighting the HST-estimated spectral radiance by the spectral response function (Fig. 4.1) of the given THEMIS-VIS band.

For bands 1 – 4, the comparison in Fig. 4.17 provides satisfactory validation of our calibration results. There are clearly systematic differences between the HST and THEMIS-VIS radiances, as well as substantial scatter, but because we are extrapolating the photometric behavior of Mars out to phase angles at which it is not adequately measured by our concurrent HST study, and because of the large difference in resolution between HST and THEMIS-VIS, the typical mismatches of 15% cannot be construed as evidence of a calibration inaccuracy. In fact, it may be possible in the future to constrain the photometric properties of Mars using this kind of comparison.

For band 5, the pre-flight-calibration-based y coefficient produced radiance values that were much too high relative to the HST radiance estimates. We have no

explanation for this, but given the very high level of stray light in band 5, it is not entirely surprising that our pre-flight calibration strategy did not produce a useful result. Table 4.2 lists our arbitrarily adopted band 5 y coefficient of 0.6 $\pm$ 0.2 DN msec^{-1}/ (W m^{-2} μm^{-1} sr^{-1}), which was selected because it produces the reasonable agreement with HST shown in Fig. 4.17.

4.6 Error Analysis

Figures 4.18 and 4.19 illustrate the significance of the two stray light contributions to the C-ROI signal in a THEMIS-VIS EDR, while Figs. 4.20 and 4.21 illustrate the way that these contributions are distributed over the area of an EDR framelet. Together with the confidence intervals on the x, y, and z calibration coefficients given in Table 4.2, these figures allow estimation of the uncertainty in any given RDR framelet's radiance values. The x and y confidence intervals enclose 95% of the one-dimensional probability density, and so can be regarded as 2σ error estimates. The z confidence interval is more problematic, since it is simply based on the difference between the two most plausible models, but in the following discussion we will treat it too as a 2σ error estimate.

For Figs. 4.18 and 4.19 we have extracted the register stray light and photosite stray light C-ROI signals, respectively, from the calibration pipeline, and compared them to the direct C-ROI signal derived by the calibration pipeline. The contribution of each type of stray light to the fractional uncertainty for a particular C-ROI radiance can be estimated by multiplying the expected fractional contribution of that type of stray light to the total signal by the fractional uncertainty in the relevant calibration coefficient. Consider, for example, a spatial summing 2, band 1 (423 nm) image with a typical perihelion-period exposure time of 2.5 millisec-

Table 4.8: Minnaert Photometric Model Coefficients

Phase angle, h, range	Wavelength		
	502 nm	672 nm	952 nm
	κ		
0° – 10°	0.727	0.778	0.859
10° – 20°	0.760	0.869	0.951
20° – 30°	0.627	0.788	0.849
30° +	0.763	0.929	1.005
	β		
0° – 10°	0.115	0.289	0.330
10° – 20°	0.103	0.283	0.345
20° – 30°	0.081	0.243	0.263
30° +	0.081	0.249	0.276

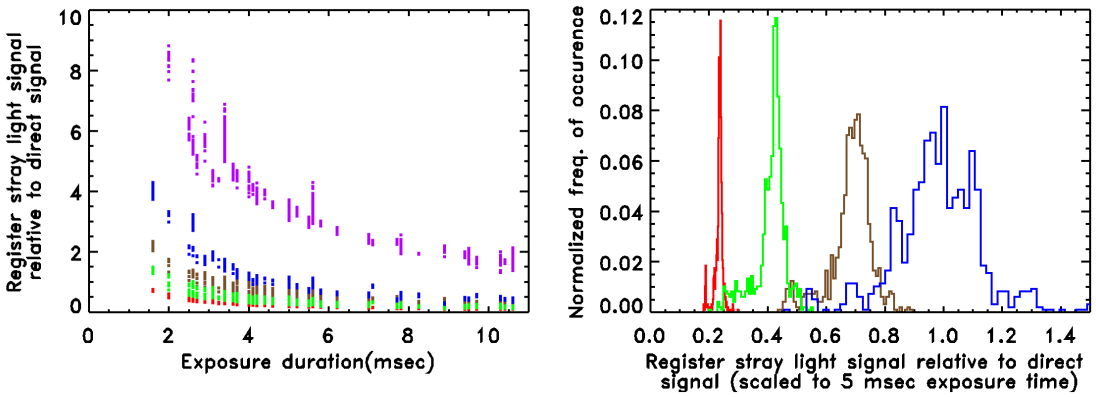

Figure 4.18: C-ROI register stray light signal as determined by the calibration pipeline for the HST-concurrent data set. The left hand panel shows the relative register signal as a function of effective exposure time for all five bands, using the same color coding as in Fig. 4.17 (blue, green, red, brown, and purple correspond to the 425, 540, 654, 749, and 860 nm bands, respectively). The right hand panel shows histograms of the same data points after the 1 / effective exposure time scaling factor has been removed. The effective exposure time is the actual exposure time multiplied by the spatial summing mode. The histograms are normalized so that the area under the histogram curve is equal to one. (Note that the histogram for the 860 nm band falls entirely beyond the range selected for the histogram plot).

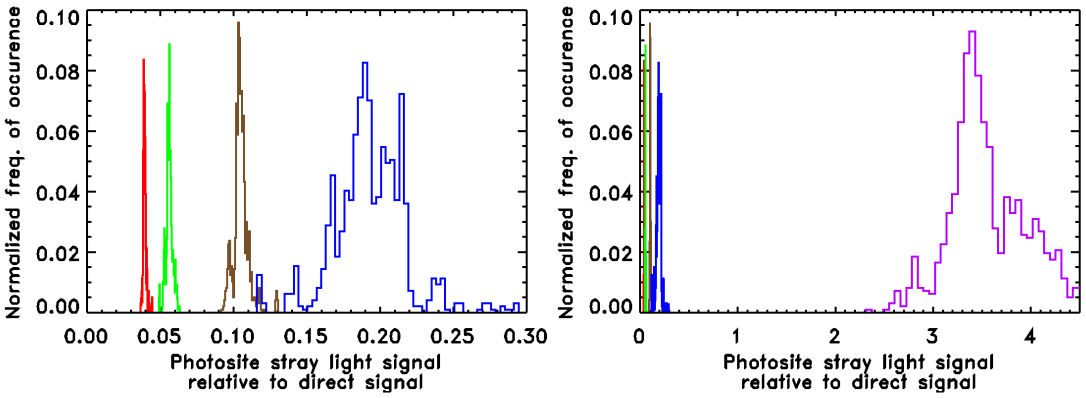

Figure 4.19: Histogram of the C-ROI photosite stray light signal as determined by the calibration pipeline, presented as a fraction of the direct C-ROI signal determined by the calibration pipeline, for the HST-concurrent data set. Color coding is the same as in Fig. 4.17 (blue, green, red, brown, and purple correspond to the 425, 540, 654, 749, and 860 nm bands, respectively). The two panels are identical, except for differing x-axis ranges. The histograms are normalized so that the area under the histogram curve is equal to one.

onds. This is an effective exposure time of 5 milliseconds ("effective" exposure time = spatial summing mode $\times$ actual exposure time), and so according to the histogram in Fig. 4.18, the register stray light signal is typically *equal* to the direct signal. The z calibration coefficient for summing mode 2 has an uncertainty of 6.4%, and so 6.4% $\times$ 1.0 gives a register stray light contribution to the uncertainty of 6.4%. For the same band and summing mode, the photosite stray light signal is typically 19%, and the x coefficient uncertainty is 8.3%, so 8.3% $\times$ 0.19 give a photosite stray light contribution to the uncertainty of 1.6%. For comparison, the uncertainty in y contributes 3.5%, and the overall 2σ uncertainty can be regarded as $\sqrt{6.4^2 + 1.6^2 + 3.5^2} = 7.5\%$, meaning that register stray light is the dominant source of error for band 1 at this particular exposure time. Table 4.9 shows similar calculations for all bands and summing modes at various exposure times.

It is important to note that since the stray light signals are additive, the residual

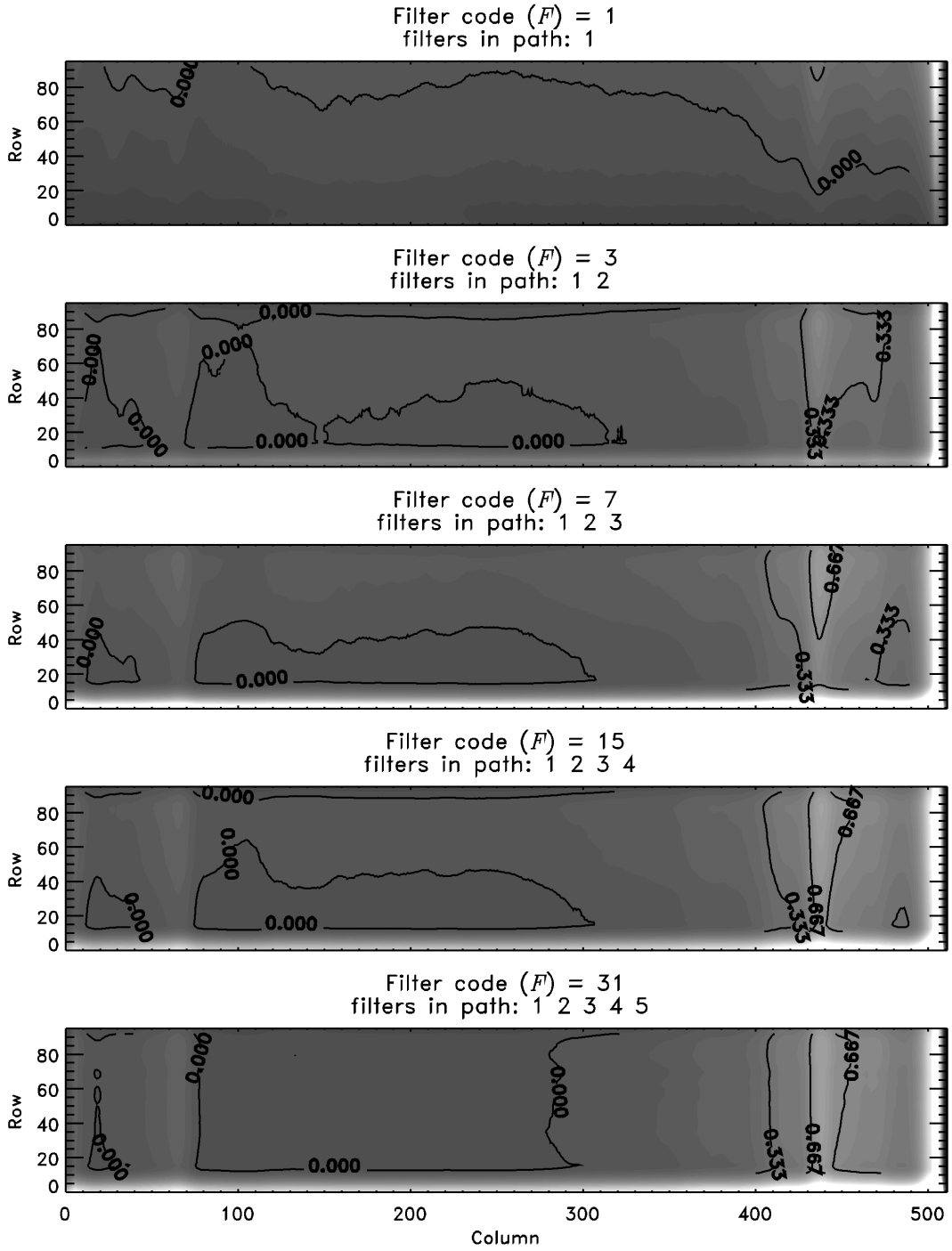

Figure 4.20: Contour plots of the pixel-dependent component of register stray light, for use in estimating errors in the calibrated C-ROI radiance based on the magnitude of visible artifacts, as discussed in the text. We show this pixel dependent component for a subset of the 31 summing mode 2 register stray light calibration frames. The values plotted are the difference between the register stray light frame and its C-ROI mean, as a fraction of it's C-ROI mean.

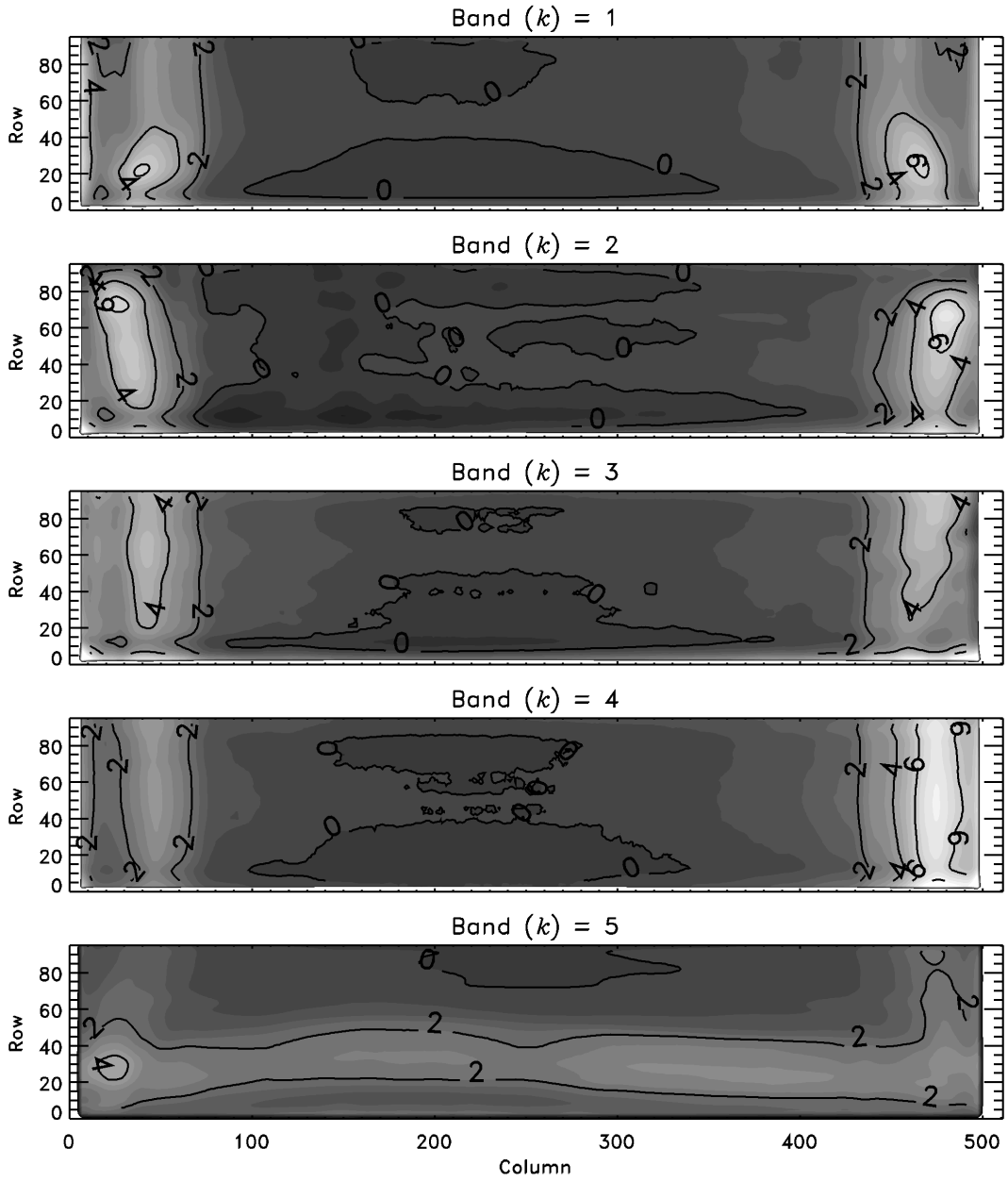

Figure 4.21: Contour plots of the pixel-dependent component of photosite stray light, for use in estimating errors in the calibrated C-ROI radiance based on the magnitude of visible artifacts, as discussed in the text. We show all 5 summing mode 2 photosite stray light frames. The values plotted are the photosite stray light frame as a fraction of the photosite stray light response coefficient, x.

Table 4.9: 2σ uncertainty as a function of band, summing mode, and effective exposure time

Band	y coefficient contribution	Photosite stray light contribution
425 nm	3.5%	1.6%
540 nm	1.2%	0.5%
654 nm	1.6%	0.3%
749 nm	2.8%	0.9%
860 nm	33.3%	53.2%

Effective Exposure Time	Band	Summing Mode 1		Summing Mode 2		Summing Mode 4	
		Register stray light contribution	Total uncertainty	Register stray light contribution	Total uncertainty	Register stray light contribution	Total uncertainty
2 msec	425 nm	17.6%	18.0%	16.1%	16.5%	36.2%	36.4%
	540 nm	7.5%	7.6%	6.9%	7.0%	15.5%	15.6%
	654 nm	4.3%	4.6%	3.9%	4.3%	8.9%	9.0%
	749 nm	12.5%	12.9%	11.5%	11.9%	25.9%	26.0%
	860 nm	61.6%	88.0%	56.5%	84.5%	127.1%	141.8%
5 msec	425 nm	7.0%	8.0%	6.4%	7.5%	14.5%	15.0%
	540 nm	3.0%	3.3%	2.8%	3.1%	6.2%	6.3%
	654 nm	1.7%	2.4%	1.6%	2.3%	3.5%	3.9%
	749 nm	5.0%	5.8%	4.6%	5.5%	10.3%	10.8%
	860 nm	24.6%	67.5%	22.6%	66.8%	50.8%	80.8%
10 msec	425 nm	3.5%	5.2%	3.2%	5.0%	7.2%	8.2%
	540 nm	1.5%	2.0%	1.4%	1.9%	3.1%	3.4%
	654 nm	0.9%	1.9%	0.8%	1.8%	1.8%	2.4%
	749 nm	2.5%	3.9%	2.3%	3.8%	5.2%	6.0%
	860 nm	12.3%	64.0%	11.3%	63.8%	25.4%	67.8%
50 msec	425 nm	0.7%	3.9%	0.6%	3.9%	1.4%	4.1%
	540 nm	0.3%	1.4%	0.3%	1.3%	0.6%	1.5%
	654 nm	0.2%	1.6%	0.2%	1.6%	0.4%	1.7%
	749 nm	0.5%	3.0%	0.5%	3.0%	1.0%	3.1%
	860 nm	2.5%	62.9%	2.3%	62.9%	5.1%	63.0%

(*i.e.*, uncorrected portion) stray light signal included in our error estimates is also additive. Thus, uncertainty introduced by residual stray light will affect the signal ratio between two regions of an image, but, in instances where we are prepared to regard residual stray light as uniform over a framelet, its uncertainty does not affect the signal difference between two regions of a framelet. Of course, uncertainty in the direct response coefficient y will affect differences, but not ratios.

Strictly speaking, all three of the sources of uncertainty (photosite stray light, register stray light, direct response coefficients) discussed thus far represent the possibility of global (*i.e.*, over the whole data set) biases. However, the success of the stray light removal for any given THEMIS-VIS image is inherently stochastic because the stray light signal depends on the brightness of the scene in regions outside of the detector field of view. The uncertainties in the global stray light calibration coefficients reflect the extent to which the stray light is unpredictable. Thus the fractional uncertainties derived from the uncertainties in the calibration coefficients can be construed as applicable to individual images as well as to the data set as a whole. Any future quantitative work with the THEMIS-VIS data set must consider stray light errors that may be both correlated and uncorrelated across the data set.

When considering regions of a framelet other than the C-ROI, the contour plots of Figs. 4.20 and 4.21 provide a scaling factor for the amount of stray light uncertainty in a calibrated RDR. These contour plots are derived from the register stray light and photosite stray light (respectively) calibration frames and coefficients. For each filter path, or each band, they are constructed by subtracting the C-ROI mean stray light signal from the calibration frames, and then dividing by the C-ROI mean stray light signal. Thus, the contour plots show the amount of stray

light contamination in a region over and above that in the C-ROI, as a fraction of the C-ROI contamination. For example, when the contour plot shows a value of 6, the amount of stray light is $6 + 1 = 7$ times as great as in the C-ROI, and the uncertainty introduced in that region by that type of stray light is 7 times as large as it is for the C-ROI. Clearly, the amount of photosite stray light increases dramatically at the edges of the framelets, whereas the increase in register stray light towards the edges of the framelets is still significant, but much less.

For brevity, we have shown these contour plots for summing mode 2 only, and for only a representative selection of filter paths. At the level of precision necessary for error estimation, these plots can nevertheless be used for any summing mode, with appropriate scaling of the axes. Similarly, they can be used for any filter path, in which case the panel of Fig. 4.20 with the same number of filters in the path as the framelet under consideration should be used. For example, the register stray light in a framelet with filters 2, 4, and 5 in its path is most similar to the values in the panel for filter 1, 2, and 3.

When calibrated (RDR) framelets can be inspected individually, and residual stray light artifacts are recognizable, Figs. 4.20 and 4.21 can also be used to provide additional information about stray light residuals. If, for example, we observe a stray light artifact in a band 1 framelet with a shape similar to the stray light lobes on the edges of the band 1 contour plot in Fig. 4.21, and we estimate the amplitude of the artifact to be 100 in some arbitrary radiance units, then since the peak of these lobes is about 7 in Fig. 4.21, the residual photosite stray light in the C-ROI is likely no more than $100/7 = 14$ of the same radiance units. Since estimating the "amplitude" of the artifact is inherently imprecise, and since stray light artifacts frequently vary dramatically in magnitude and sign across a framelet

and between neighboring framelets, this type of calculation should normally be treated as a constraint on the error rather than as a correction to the RDR radiance. It may ultimately be possible to derive corrections to RDR radiances by iteratively adjusting, scaling, and subtracting calibration frames until stray light artifacts are no longer visible, but that work is beyond the scope of this chapter.

Most calibrated THEMIS-VIS images show stray light artifacts to some degree, although, as Table 4.9 makes clear, these artifacts are more significant in band 1, and least significant in band 3. In band 5, the residual stray light artifacts normally dominate the direct signal, so that band 5 only very rarely provides usable data. In Figs. 4.22 and 4.23 we present two examples of THEMIS-VIS RDRs with very prominent stray light artifacts. Figure 4.22 is dominated by register stray light residuals, whereas Fig. 4.23 shows a combination of both effects. The most diagnostic characteristic of register stray light is the pattern of narrow stripes that it produces at the edges of each framelet. As is often the case, because it is far from the readout register, and its direct signal is relatively low, band 4 shows the most prominent pattern of narrow stripes. However, as is apparent from Fig. 4.9 and Fig. 4.20, the narrow stripes only become a significant part of the register stray light pattern for rows far from the h-register, and so given that band 4 also shows a uniformly positive register stray light residual, it is probable that band 1 also has a significant register stray light residual even though band 1 shows no obvious pattern. Another symptom of register stray light is a "roll-off" (or "roll-up") near the top or bottom of each framelet caused by the interaction of the flatfield with the register stray light subtraction. This roll-off often appears as a bright or dark stripe. When such a roll-off is present, it is most likely caused by register stray light, simply because it is the most important source of error. Photosite stray light

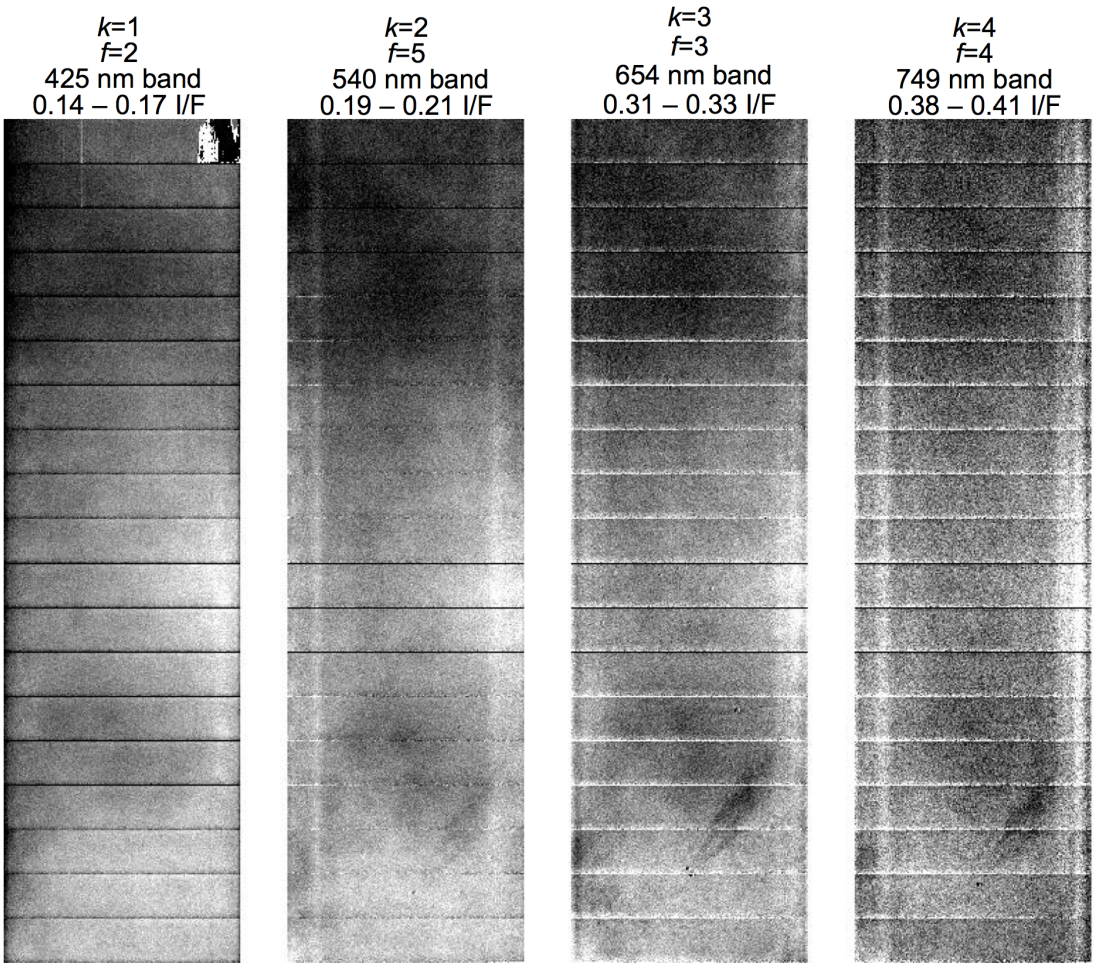

Figure 4.22: An example RDR with severe register stray light residuals. (Image number V12386007)

residuals may also lead to a roll-off, however.

Unlike Fig. 4.22, where the residuals are clearly one particular kind of stray light, and are quite uniform throughout the image, Fig. 4.23 shows highly variable residuals and a combination of register and photosite stray light. The abrupt changes in brightness between framelets could be caused by changes in either photosite stray light or register stray light residuals, but register stray light is always the most likely culprit for ambiguous artifacts. However, some features in the

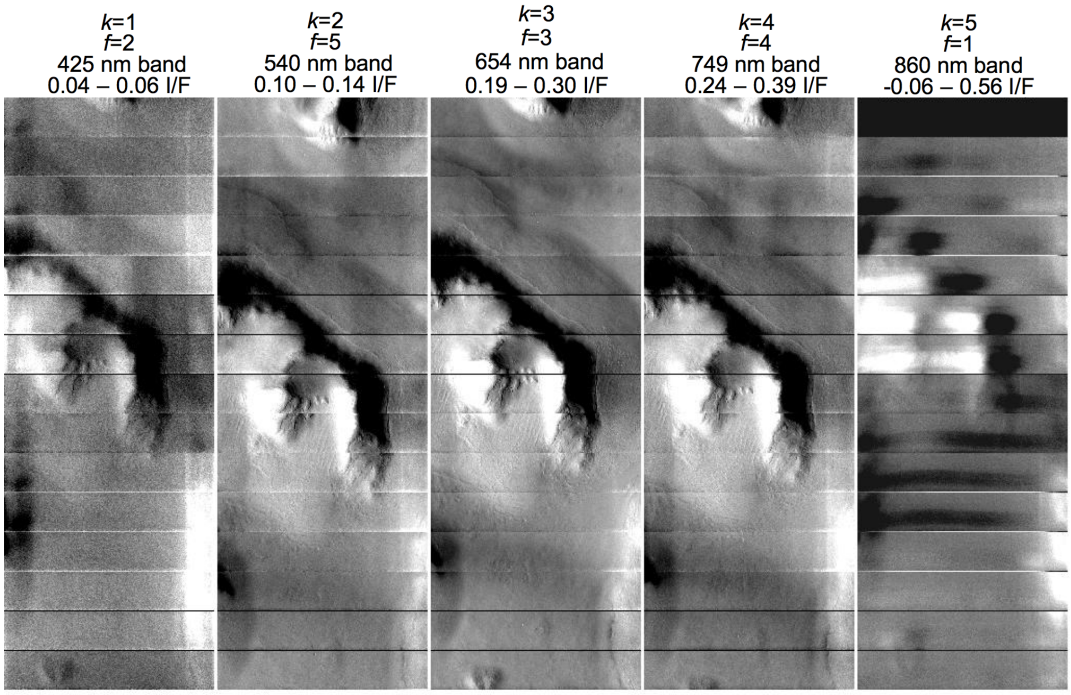

Figure 4.23: An example RDR with severe photosite stray light and register stray light residuals. (Image number V02699003)

Fig. 4.23 RDR are readily distinguishable as photosite stray light, because residual photosite stray light produces a broad swath or lobe at the framelet edges for bands 1 – 4 rather than a narrow stripe. Of course, these lobes can be either bright or dark depending on the sign of the residual. The bottom six framelets of band $k = 1$ show the clearest example of these lobes, with a broad swath of high positive residual on the right hand side accompanied by negative residuals on the left hand side. Presumably this asymmetry is caused by bright and dark scene elements just outside the detector's field of view. In band 5 the broad curving photosite stray light stripe is almost always clearly visible as a residual, which appears to be modulated by the brightness of the scene just below the framelet's field of view.

4.7 Comparison with other data sets

We conclude by evaluating the extent to which I/F values based on our calibrated radiances are comparable with historical HST measurements of Mars, and with Mars Exploration Rover Pancam measurements from the Martian surface. The solar irradiance used to calculate I/F for a given THEMIS VIS band is obtained by taking a weighted average of the solar irradiance spectrum, where the weights are that band's relative response function (Figure 4.1a). Since the historical HST data and the Pancam data were acquired at different times and with different viewing geometries from any THEMIS-VIS images we might compare them to, we do not expect them to match in terms of absolute radiances. Instead, we will compare normalized I/F spectra in order to demonstrate that the relative spectral calibration of THEMIS-VIS is reasonable, and to demonstrate some of THEMIS-VIS's scientific potential.

To minimize the impact of atmospheric variability on the HST comparison, we

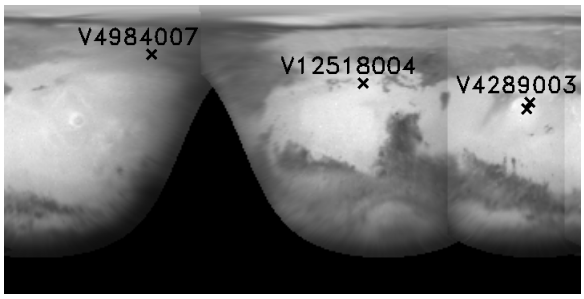

Figure 4.24: Simple cylindrical, 1 pixel-per-degree projection of HST 763 nm imaging from March 3, 1997 (see Table 4.3 for details), showing the location of HST pixels used for comparison with THEMIS-VIS spectra in Figs. 4.26, 4.28, and reffig30. The HST data used for these comparisons are not photometrically corrected, which leaves visible seams when the images are mosaiced together.

have selected a set of HST observations from northern summer (March 10, 1997, L_s 89°, see Table 4.3 for details), and chosen our THEMIS-VIS images from mid-northern latitudes in the 90° - 120° L_s period. Consulting Smith (2004), we note that this is Mars's most consistently aerosol-free region and season. Figure 4.24 shows the March 1997 HST data and the locations of the HST data points which we compare to the THEMIS-VIS data. Each location is labeled with the number of its corresponding THEMIS-VIS image.

For the first comparison, Fig. 4.25 and 4.26, we have selected two HST data points in the vicinity of Elysium Mons, one inside of (at 25° N) and one outside of (at 29° N) the cloud feature on the volcano's flank. The THEMIS-VIS image (Fig. 4.25) also shows a clear region to the north and water-ice cloud in the south, and so we plot spectra (Fig. 4.26) from both areas. Since the THEMIS-VIS and HST spectra are from different times of day ($\sim$ 5pm local solar time for THEMIS and near the sub-solar point for HST), slightly different locations relative to a variable cloud feature, and two martian years apart in time, it is not surprising that the spectral match is not exact. However, the difference in normalized I/F caused

by the cloud is very similar in the HST and the THEMIS-VIS data, suggesting THEMIS-VIS is correctly measuring cloud-induced spectral contrasts. This validative evidence is important, because THEMIS-VIS spectral data is frequently notable for its lack of spectral contrast see, for example, Figs. 4.29 and 4.30 or Figs. 4.31 and 4.32. By showing the correct degree of spectral contrast in Fig. 4.26, we demonstrate that the spectral uniformity on THEMIS-VIS scales is in many locations a real characteristic of the martian surface.

Figures 4.27 and 4.28 show that without the confounding influence of aerosols, the THEMIS-VIS data very closely match the historical HST results. For the brighter, redder terrain unit, which predominates in this region of northern Arabia Terra and thus controls the much lower resolution HST spectrum, the THEMIS-VIS spectrum is a very close match. In Figure 4.30, however, we don't have a good match between the HST and THEMIS-VIS spectra. We do not have an explanation for this mismatch, but the combination of variable atmospheric hazes with high atmospheric path length in the HST data (emission angle of 49°), and changes in surface dust cover, are both possibilities.

Figures 4.31 – 4.34 compare THEMIS-VIS spectra with published Pancam data [Bell et al., 2004a,c]. In Figs. 4.31 – 4.32 we select examples of the two main surface albedo units in the vicinity of the Spirit Rover landing site, avoiding the region of excess residual stray light at the bottom of the THEMIS-VIS image (Fig. 4.31), and find that these surface units have a distinct but very small color contrast when viewed from orbit. The THEMIS-VIS spectra are slightly bluer than the Pancam soil measurement, which could be explained by the extra atmospheric path length in the THEMIS-VIS observations, or by areal mixing with the sparse population of exposed rock.

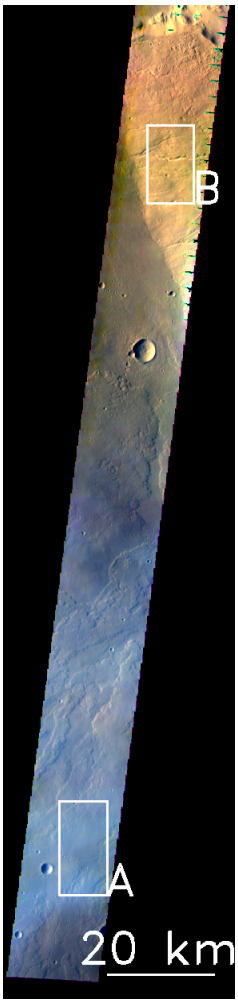

Figure 4.25: Sinusoidal projection of THEMIS-VIS image V04289003 with exaggerated colors. North is at the top, and east is at right. The band 1, 2, and 3 images are used for the blue, green, and red channels, respectively, with each channel stretched arbitrarily to maximize contrast. Spectra from regions A and B are shown in Fig. 4.26. These regions are centered at: A: 29.54° N, 148.48° E, incidence angle 64°; B: 31.69° N, 148.80° E, incidence angle 64°.

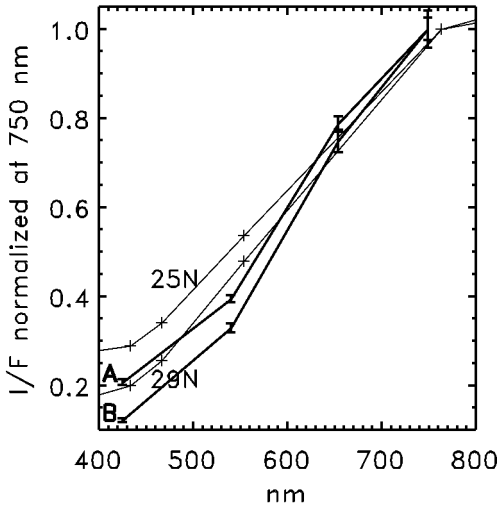

Figure 4.26: HST WFPC2 and THEMIS-VIS (thick lines) I/F normalized to 1 at ∼750nm. The THEMIS-VIS spectra are from V04289003, and their labels correspond to the A and B regions in Fig. 4.25. The HST spectra, labeled by their latitudes 29 N and 25 N, correspond to the two points labeled with "V04289003" in Fig. 4.24. The THEMIS-VIS error bars show the standard deviation of I/F in the region from which the spectra were derived.

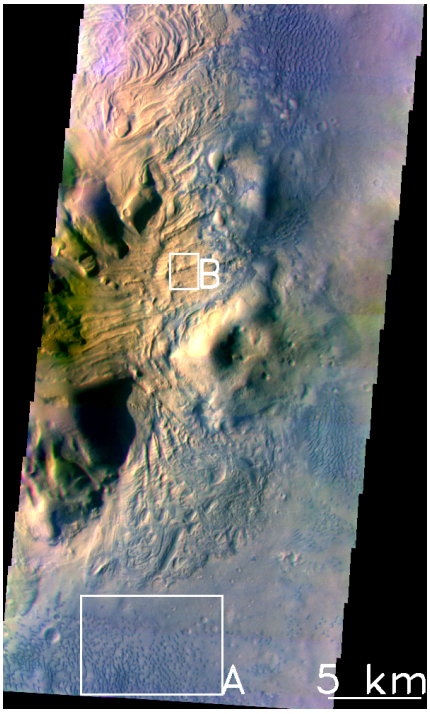

Figure 4.27: Portion of THEMIS-VIS image V12518004, with sinusoidal projection and exaggerated colors (using bands 1, 2, and 3) as in Fig. 4.25. Spectra from regions A and B are shown in Fig. 4.28. These regions are centered at: A: 41.44° N, 44.66° E, incidence angle 68°; B: 41.78° N, 44.70° E, incidence angle 68°.

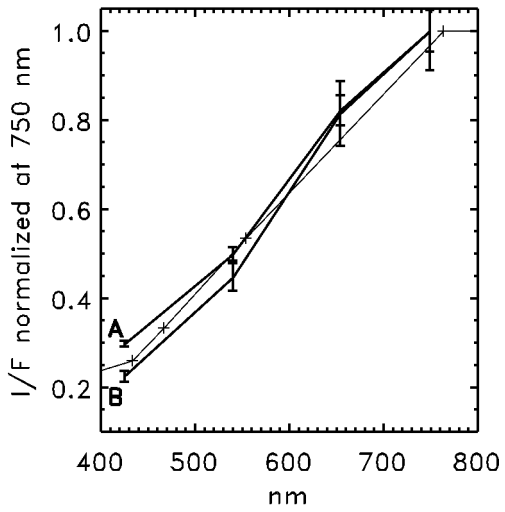

Figure 4.28: HST WFPC2 and THEMIS-VIS (thick lines) I/F normalized to 1 at ∼750nm. The THEMIS-VIS spectra are from V12518004, and their labels correspond to the A and B regions in Fig. 4.27. The HST spectrum is taken from the point labeled "V12518004" in Fig. 4.24. The THEMIS-VIS error bars show the standard deviation of I/F in the region from which the spectra were derived.

THEMIS-VIS V03671001 in Fig. 4.33 does not quite cover the Opportunity Rover landing site, but it is the only multi-band image of the landing site vicinity without severe atmospheric aerosol contamination. As in Figs. 4.31 – 4.32, our spectra sample two distinct albedo/color units, and avoid the residual stray light. Pancam clearly shows much greater spectral diversity than THEMIS-VIS, but the THEMIS-VIS spectra fall well within the range of the Pancam samples. The dark terrain THEMIS-VIS spectrum ("A") is a good match to the dark Eagle Crater soil ("5"), and the lighter THEMIS-VIS terrain ("B") is redder just as the bright Pancam outcrop ("1") is, but to a much lesser degree.

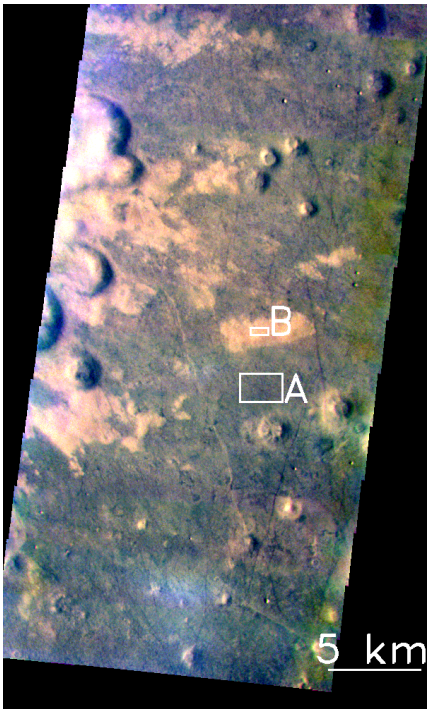

Figure 4.29: Portion of THEMIS-VIS image V04984007, with sinusoidal projection and exaggerated colors (using bands 1, 2, and 3) as in Fig.4.25. Spectra from regions A and B are shown in Fig. 4.30. These regions are centered at: A: 60.02° N, 272.73° E, incidence angle 69°; B: 60.07° N, 272.72° E, incidence angle 69°.

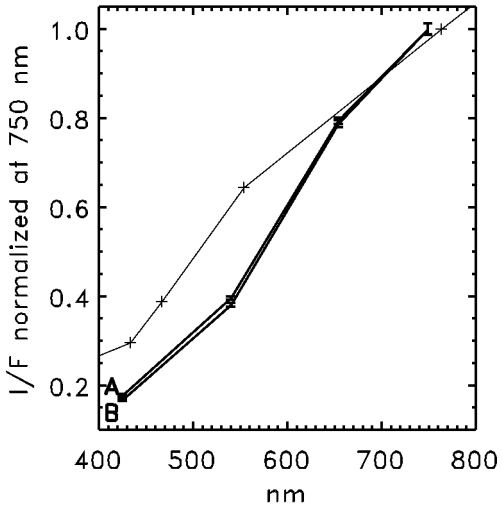

Figure 4.30: HST WFPC2 and THEMIS-VIS (thick lines) I/F normalized to 1 at ∼750nm. The THEMIS-VIS spectra are from V04984007, and their labels correspond to the A and B regions in Fig. 4.29. The HST spectrum is taken from the point labeled "V04984007" in Fig. 4.24. The THEMIS-VIS error bars show the standard deviation of I/F in the region from which the spectra were derived.

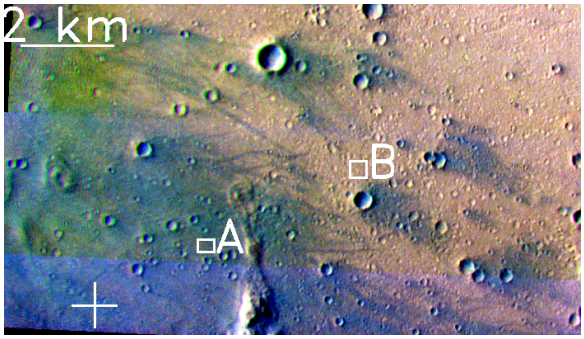

Figure 4.31: Portion of THEMIS-VIS image V10792003, with sinusoidal projection and exaggerated colors (using bands 1, 2, and 3) as in Fig. 4.25. Spectra from regions A and B are shown in Fig. 4.32. These regions are centered at: A: 14.548° S, 175.514° E, incidence angle 75°; B: 14.520° S, 175.571° E, incidence angle 75°. The cross marks the MER-Spirit landing site at 14.5692° S, 175.4729° E [Squyres et al., 2004a].

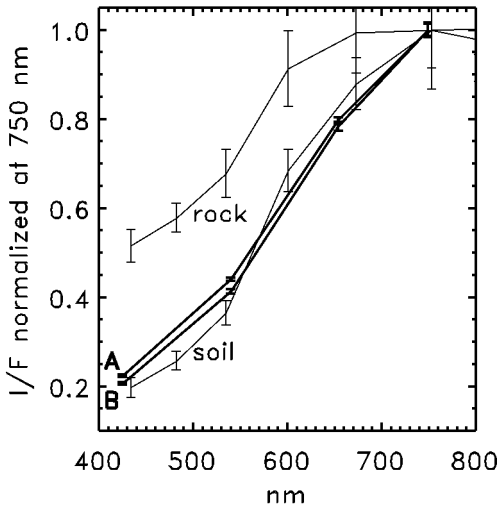

Figure 4.32: MER-Pancam and THEMIS-VIS (thick lines) I/F normalized to 1 at ∼750 nm. The THEMIS-VIS spectra are from V10792003, and their labels correspond to the A and B regions in Fig. 4.31. The Pancam spectra show a soil and a rock surface near the Spirit landing site, and are taken from Fig. 6C in Bell et al. [2004a]. For both data sets, the error bars show the standard deviation of I/F in the region of the image from which that spectrum was derived.

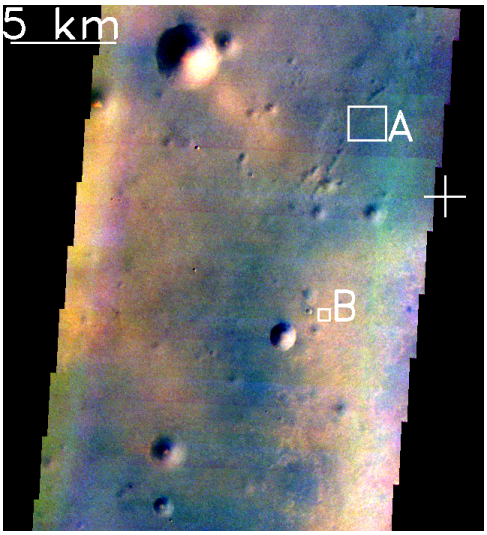

Figure 4.33: Portion of THEMIS-VIS image V03671001, with sinusoidal projection and exaggerated colors (using bands 1, 2, and 3) as in Fig. 4.25. Spectra from regions A and B are shown in Fig. 4.34. These regions are centered at: A: 1.888° S, 354.405° E, incidence angle 71°; B: 2.040° S, 354.370° E, incidence angle 71°. The cross marks the MER-Opportunity landing site at 1.9462° S, 354.4734° E [Squyres et al., 2004b].

4.8 Conclusions

Despite significant stray light contamination, we have obtained an absolute calibration for THEMIS-VIS 654 nm bandpass images with an uncertainty that we estimate to be better than 5% (2σ) in most situations. The calibration uncertainty is greater for the 425 nm, 540 nm, and 749 nm bands, but even for these bands the uncertainty (Table 4.9) can be better than 5% (2σ) in low-light conditions where long exposure times are possible. For the 860 nm band, however, the uncertainties in the radiances are so high that they are effectively un-usable. For all of the THEMIS-VIS bands, relative radiance comparisons are no more precise than absolute measurements, because the additive and variable residual stray light is responsible for most of the uncertainty.

Generally speaking, the exposure duration for an imaging sequence is limited

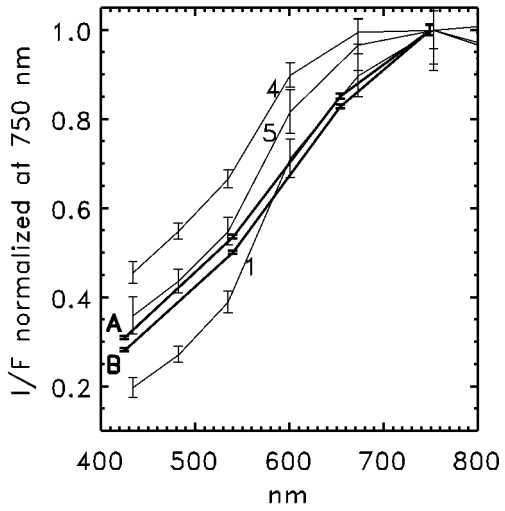

Figure 4.34: MER-Pancam and THEMIS-VIS (thick lines) I/F normalized to 1 at ∼750 nm. The THEMIS-VIS spectra are from V03671001, and their labels correspond to the A and B regions in Fig. 4.33. The Pancam spectra labeled 1, 4, and 5 correspond to the labels in Fig. 8A of Bell et al. [2004c]. These Pancam spectra show a bright outcrop (1), dark plains (4), and the Eagle crater soil (5). For both data sets, the error bars show the standard deviation of I/F in the region of the image from which that spectrum was derived.

by saturation in the most sensitive bands (654 nm and 540 nm, in that order). Thus, since all bands in a sequence have the same exposure duration, leaving the most sensitive bands out of a sequence can reduce the contribution of register stray light in the remaining bands, significantly improving their uncertainty over bright surfaces. Of course, sacrificing the more sensitive bands for this purpose is only rarely desirable.

THEMIS-VIS spectra are consistent with Pancam measurements, and are mostly consistent with historical HST measurements. Furthermore, our concurrent-imaging calibration validation strategy indicates that THEMIS-VIS absolute radiances and HST absolute radiances are broadly consistent. However, the concurrent imaging does not in and of itself provide a strong constraint on the accuracy of THEMIS-VIS absolute radiances because of the limitations in the ability of the photometric models to "correct" the observations to the same standard viewing geometry. The comparison does suggest, however, that THEMIS-VIS could in the future play a role in constraining Mars's photometric properties.

Residual stray light artifacts can frequently be quite prominent in RDR images that have be stretched to maximize contrast, but because photosite stray light is so much more intense near the framelet edges, the presence of visible photosite stray light artifacts does not normally indicate a significant problem with center-field radiances. Obvious register stray light artifacts are more rare, but they are a more serious concern for the accuracy of center-field radiances. When quantitative radiance values are important, THEMIS-VIS data users should check their accuracy by comparing any prominent artifacts with Figs. 4.20 and 4.21 of this chapter.

Notation

a	exposure number (0 – . . .)
B	bias frame, *i.e.*, bias charge for a framelet (DN)
C-ROI	region of interest for which calibration coefficients are defined
$\overline{d}$	difference in mean raw DN values for a pair of framelets
D	raw (bias subtracted, for cases where it matters) 11-bit DN value
$\overline{D}$	C-ROI mean raw 11-bit DN value
E	characteristic 'bias' charge from the portion of the chip under a certain filter (DN)
f	filter number (1 – 5)
F	filter path code (1 – 31)
$\overline{g}$	difference in the mean normalized register stray light frames of a pair of framelets
G	normalized register stray light frame, *i.e.*, spatial pattern of register stray light for a framelet (unitless)
$\overline{G}$	C-ROI mean of a normalized register stray light frame
h	phase angle
i	column number (0 – . . .)
I_k	radiance in a particular band
$\overline{I}_k$	C-ROI mean radiance in a particular band
$\overline{I}'_k$	deviations from the mean radiance
$\widehat{\overline{I}}$	C-ROI mean broadband radiance
j	row number (0 – . . .)
k	band number (1 – 5)
l	ground calibration lamp level index
m	framelet number (0 – . . .)
n	number of framelets in an EDR or RDR

N	number of elements in the set of framelets used to estimate R
p	an index for the set of all framelets in the data set
q	an index for a set of framelets or data points used in various contexts
Q	calibrated THEMIS-VIS signal (DN/msec)
R	responsivity (*i.e.*, the flatfield) (unitless)
$\overline{S}$	C-ROI mean photosite signal (DN/msec)
S	photosite signal (DN/msec)
S^*	ground calibration photosite signal prior to background subtraction (DN/msec)
t	exposure duration
u	index for the set of framelet pairs used to generate the register stray light response coefficient
u'	index for the set of framelet pair members
v	elements of the sample of framelets used to estimate ψ
x_k	photosite stray light response coefficient
X	photosite stray light calibration frame, *i.e.*, the spatial pattern of photosite stray light for a framelet
y_k	direct response coefficient
z	register stray light response coefficient
Z	register stray light (DN) frame in ground calibration
α_k	coefficients for predicting $\overline{\widehat{I}}$ from $\overline{I}_k$
β	Minnaert photometric model coefficient

CHAPTER 5

SUMMARY AND FUTURE WORK

In this work we have described the behavior of the martian polar vortices over a period of three martian years, using MGS-TES derived Ertel potential vorticity (*PV*) on isentropic surfaces as our primary analysis quantity; we have confirmed that the martian equatorial mesosphere produces discrete clouds; and we have presented the first evidence for discrete clouds in the winter mid-latitude mesosphere. We have also presented and evaluated the standard calibration pipeline for the THEMIS-VIS imager that we use to detect and measure the mesospheric clouds.

On Mars, as on Earth, the buildup of *PV* in the winter polar region is dominated by radiative cooling. However, whereas the terrestrial stratospheric polar vortices are most of the time defined by a single high-*PV* airmass with a *PV* field that, although highly asymmetric, increases monotonically towards the pole, we find that the martian polar vortices are characterized by an annular *PV* field, which is relatively symmetric but reaches a maximum between 60° and 80° latitude and then decreases towards the pole. This change in sign of the vorticity gradient raises the possibility of barotropic instability on the poleward flank of the winter polar jet. However, the applicability of the standard quasi-geostrophic barotropic instability criterion to the Rossby number ~ 1 flows of the martian polar vortices will have to be investigated.

The northern polar vortex is substantially more intense and better organized than the southern polar vortex, suggesting that the northern vortex is a stronger barrier to mixing and heat transport. The northern vortex *PV* maximum is observed to move poleward in response to dust loading in the low-latitude atmosphere, and the onset of dust storms is associated with large transient wave perturbations of

the vortex. In one case a dust storm appears to trigger a wave-breaking event that is connected with an abrupt warming in the core of the vortex.

The southern vortex exhibits a great deal of variability that is not associated with any obvious external factors. The most dramatic event we observe in the southern vortex is a 10 K warming event associated with a partial breakdown of the *PV* field and a reversal in wind direction.

The polar vortex study that we have presented is primarily descriptive, with the intent of providing a basis for theoretical and modeling studies that will help clarify the role of *PV* and the polar vortex in isolating the winter polar atmosphere. The most basic need is a scheme for estimating radiative heating/cooling rates from the TES data set, which will enable us to evaluate Q at all of our grid points and therefore understand in detail where and how *PV* is being created and destroyed, and to improve on our order-of-magnitude estimates of the rates at which this occurs. A complete picture of Q would also require an estimate of CO_2 condensation rates.

Equally important are comparisons with GCMs. We need to evaluate the extend to which GCMs match the annular *PV* configuration, and the rapid response to dust loading, and whether their scenarios for sudden polar warmings rearrange the *PV* field in ways consistent with our observations.

Another way to evaluate the isolating properties of the polar vortex is to use dust aerosol as a tracer, observing whether or not dust injected into the atmosphere by dust storms penetrates to the pole. To do so, we must develop techniques to derive dust opacities from limb-pointed TES measurements. Standard nadir-pointed TES dust measurements are not valid over the winter poles because low surface temperatures lead to poor contrast in dust absorption features.

In order to perform our *PV* analysis of the polar vortices, we have developed techniques to interpolate MGS-TES temperature retrievals onto a regular grid, and to solve for a horizontal wind field with much greater detail than previous work. This data set has a wide range of potential applications. One obvious example is comparisons with imaging studies of cloud motions such as those of Wang and Ingersoll [2002] and Wang et al. [2005]. A more ambitious application of our regularly gridded temperatures is principal components analysis to elucidate the normal modes of variability in the polar atmosphere, a technique which has been used to study the terrestrial stratospheric annular modes. [e. g. Thompson et al., 2000]

Although our gridded winds fields are presently limited by the MGS-TES nadir-pointed retrievals to the lower 3.5 pressure scale heights of the atmosphere, incorporating the limb-sounding MGS-TES temperatures will allow us to extend it to mesospheric altitudes and allow comparison with the direct wind measurements from THEMIS-VIS. The THEMIS-VIS velocities might also be incorporated as a boundary condition for the balance wind solution, which may be preferable to the arbitrary zero-wind level that we currently enforce. Another, speculative, point of comparison between our gridded fields and the THEMIS-VIS mesospheric clouds is the presence of symmetric instability at the top of our grid at latitudes comparable to those of the mid-latitude winter clouds. We might explore the hypothesis that some mid-latitude mesospheric clouds are associated with symmetric instability, just as the equatorial mesospheric clouds are hypothesized to be associated with breaking gravity waves.

The study of mesospheric clouds presented here is, like the polar vortex study, primarily descriptive. Confirmation of their existence is a significant result in and

of itself. The 5 examples of equatorial mesospheric clouds are concentrated in the eastern Tharsis / Valles Marineris region, and the northern mid-latitude clouds are seen only after sunset and are concentrated in the Acidalia region. Equatorial mesospheric clouds with detectable optical depth are thus probably rare, and their occurrence is probably biased towards eastern Tharsis / Valles Marineris, but firm conclusions about the distribution of these clouds will require discovery of more examples. The northern mid-latitude mesospheric clouds are much more common ($\sim 15\%$ detection rate), and may be biased towards the Acidalia region, or perhaps simply biased toward higher latitudes. Invisibility until after sunset at the local surface is probably a real feature of these clouds, but a bias towards the fall and early winter over mid- to late winter cannot be ruled out as an explanation for the observed pattern of detections.

One example of the value of mesospheric cloud detections is as tracers of mesospheric winds. We have demonstrated this by comparing our measured cloud velocities with GCM winds. We find good agreement with the model for the mid-latitude winter clouds, and poorer agreement for the equatorial clouds. The measured winds provide a new constraint on GCMs, which will hopefully lead to model refinements such as improved parameterizations of gravity-wave drag.

Martian mesospheric clouds can also yield insights into mesospheric temperatures, cloud microphysics, and chemistry, but all of these applications require information about the nature and number of aerosol particles. We have applied a simple radiative transfer model to demonstrate that THEMIS-VIS is capable of yielding constraints on the particle properties. The fact that this is possible serves as justification for developing more complicated radiative transfer models that will enable some assumptions to be relaxed, and allow the twilight geometry of the

mid-latitude clouds to be addressed.

The mesospheric cloud study relies crucially on the THEMIS-VIS radiometric calibration which we present in the final chapter. The altitude and velocity measurements require minimal spatially variable stray light artifacts to avoid false positives and biases in the cross-correlations. The radiative transfer models require accurate radiance measurements in as many filters as possible. As shown in Table 4.9, we have achieved an adequate radiometric calibration for four of the five THEMIS-VIS filters, despite the presence of serious stray light contamination. For long exposure times, the 2σ calibration uncertainty is better than 5% for all four usable filters.

Despite the successful radiometric calibration, the spatial signature of calibration artifacts is still commonly visible in THEMIS-VIS RDRs. These calibration residuals represent one avenue for future improvement of the THEMIS-VIS calibration. Since we know the magnitude of the spatial stray light patterns relative to the center field stray light components (Figures 4.21 and 4.22), if we interactively remove the spatial calibration residual by applying linear combinations of the stray light patterns, we can calculate the corresponding correction to the center field stray light and thus correct the absolute radiance for the entire framelet. Unless the human judgement required to separate calibration artifacts from real image features can somehow be automated, this interactive correction would be applied only to a small number of "high value" THEMIS-VIS images, such as those in which mesospheric clouds are found.

Another direction for improvement in the THEMIS-VIS calibration is to apply the simple radiative transfer models that we used on the mesospheric aerosols to the task of "removing" the atmosphere, i. e., solving for the true surface albedo

using external information about the intervening aerosols. Without atmospheric removal, comparison of surface spectral information in images acquired in different seasons or regions is confounded by variations in the contribution of atmospheric aerosols to the observed radiances. For example, typical optical depths of water ice aerosol can contribute 30% of the total observed radiance at the wavelength of the THEMIS-VIS blue filter [Clancy et al., 2003]. This problem can be solved with an application of the DISORT model that is much simpler than the one described in Chapter 3: the dust and water ice aerosol optical depths are taken from MGS-TES or THEMIS-IR measurements, standard assumptions are made about the aerosol scattering properties, and the surface albedo is iteratively adjusted until the observed radiance is matched. To make this process computationally efficient, a lookup table can be developed for surface albedo as a function of radiance, viewing geometry, and optical depth.

BIBLIOGRAPHY

Acton, C. H., 1996. Ancillary data services of NASA's Navigation and Ancillary Information Facility. Planetary and Space Sciences 44, 65–70.

Andrews, D. G., Holton, J. R., Leovy, C. B., 1987. *Middle atmosphere dynamics*. Academic Press, Orlando.

Atreya, S. K., Blamont, J. E., 1990. Stability of the Martian atmosphere - Possible role of heterogeneous chemistry. Geophys. Res. Lett. 17, 287–290.

Baldwin, M. P., Dunkerton, T. J., 2001. Stratospheric Harbingers of Anomalous Weather Regimes. Science 294, 581–584.

Bandfield, J. L., Hamilton, V. E., Christensen, P. R., McSween, H. Y., 2004. Identification of quartzofeldspathic materials on Mars. Journal of Geophysical Research (Planets) 109 (E18), 10009–+.

Banfield, D., Conrath, B., Pearl, J. C., Smith, M. D., Christensen, P., 2000. Thermal tides and stationary waves on Mars as revealed by Mars Global Surveyor thermal emission spectrometer. J. Geophys. Res. 105, 9521–9538.

Banfield, D., Conrath, B. J., Gierasch, P. J., Wilson, R. J., Smith, M. D., 2004. Traveling waves in the martian atmosphere from MGS TES Nadir data. Icarus 170, 365–403.

Banfield, D., Conrath, B. J., Smith, M. D., Christensen, P. R., Wilson, R. J., 2003. Forced waves in the martian atmosphere from MGS TES nadir data. Icarus 161, 319–345.

Barnes, J. R., Hollingsworth, J. L., 1987. Dynamical modeling of a planetary wave mechanism for a Martian polar warming. Icarus 71, 313–334.

Barnes, J. R., Pollack, J. B., Haberle, R. M., Leovy, C. B., Zurek, R. W., Lee, H., Schaeffer, J., 1993. Mars atmospheric dynamics as simulated by the NASA AMES General Circulation Model. II - Transient baroclinic eddies. J. Geophys. Res. 98, 3125–3148.

Bell, J. F., III, 2003. HST studies of Mars. In: Livio, M., Noll, K., Stiavelli, M. (Eds.), A Decade of Hubble Space Telescope Science, pp. 1–24.

Bell, J. F., III, Noe Dobrea, E. Z., Hubbard, M. Y. H., Wolff, M. J., Noll, K., Lubenow, A., Morris, R. V., Videen, G., Shkuratov, Y., 2004b. HST WFPC2, ACS, and STIS Observations of Mars During the 2003 Perihelic Opposition. Bulletin of the American Astronomical Society 36, 1182–+.

Bell, J. F., III, Squyres, S. W., Arvidson, R. E., Arneson, H. M., Bass, D., Blaney, D., Cabrol, N., Calvin, W., Farmer, J., Farrand, W. H., Goetz, W., Golombek, M., Grant, J. A., Greeley, R., Guinness, E., Hayes, A. G., Hubbard, M. Y. H., Herkenhoff, K. E., Johnson, M. J., Johnson, J. R., Joseph, J., Kinch, K. M., Lemmon, M. T., Li, R., Madsen, M. B., Maki, J. N., Malin, M., McCartney, E., McLennan, S., McSween, H. Y., Ming, D. W., Moersch, J. E., Morris, R. V., Noe Dobrea, E. Z., Parker, T. J., Proton, J., Rice, J. W., Seelos, F., Soderblom, J., Soderblom, L. A., Sohl-Dickstein, J. N., Sullivan, R. J., Wolff, M. J., Wang, A., 2004a. Pancam Multispectral Imaging Results from the Spirit Rover at Gusev Crater. Science 305, 800–807.

Bell, J. F., III, Squyres, S. W., Arvidson, R. E., Arneson, H. M., Bass, D., Calvin, W., Farrand, W. H., Goetz, W., Golombek, M., Greeley, R., Grotzinger, J., Guinness, E., Hayes, A. G., Hubbard, M. Y. H., Herkenhoff, K. E., Johnson, M. J., Johnson, J. R., Joseph, J., Kinch, K. M., Lemmon, M. T., Li, R., Madsen, M. B., Maki, J. N., Malin, M., McCartney, E., McLennan, S., McSween, H. Y., Ming, D. W., Morris, R. V., Dobrea, E. Z. N., Parker, T. J., Proton, J., Rice, J. W., Seelos, F., Soderblom, J. M., Soderblom, L. A., Sohl-Dickstein, J. N., Sullivan, R. J., Weitz, C. M., Wolff, M. J., 2004c. Pancam Multispectral Imaging Results from the Opportunity Rover at Meridiani Planum. Science 306, 1703–1709.

Bell, J. F., III, Wolff, M. J., James, P. B., Clancy, R. T., Lee, S. W., Martin, L. J., 1997. Mars surface mineralogy from Hubble Space Telescope imaging during 1994-1995: Observations, calibration, and initial results. J. Geophys. Res. 102, 9109–9124.

Benson, J. L., Bonev, B. P., James, P. B., Shan, K. J., Cantor, B. A., Caplinger, M. A., 2003. The seasonal behavior of water ice clouds in the Tharsis and Valles Marineris regions of Mars: Mars Orbiter Camera Observations. Icarus 165, 34–52.

Bohren, C. F., Huffman, D. R., 1983. Absorption and scattering of light by small particles. Technical report, University of Arizona.

Bonello, G., Bibring, J.-P., Poulet, F., Gendrin, A., Gondet, B., Langevin, Y., Fonti, S., 2004. Visible and infrared spectroscopy of minerals and mixtures with the OMEGA/MARS-EXPRESS instrument. Planetary and Space Sciences 52, 133–140.

Bridger, A. F. C., Haberle, R. M., Hollingsworth, J. L., 2003. Interannual Atmospheric Variability Simulated by a Mars GCM: Impacts on the Polar Regions. In: Third International Conference on Mars Polar Science and Exploration, pp. 8111.

Cantor, B., Malin, M., Edgett, K. S., 2002. Multiyear Mars Orbiter Camera (MOC) observations of repeated Martian weather phenomena during the northern summer season. Journal of Geophysical Research (Planets) 107, 3–1.

Chassefière, E., Blamont, J. E., Krasnopolsky, V. A., Korablev, O. I., Atreya, S. K., West, R. A., 1992. Vertical structure and size distributions of Martian aerosols from solar occultation measurements. Icarus 97, 46–69.

Christensen, P. R., Anderson, D. L., Chase, S. C., Clancy, R. T., Clark, R. N., Conrath, B. J., Kieffer, H. H., Kuzmin, R. O., Malin, M. C., Pearl, J. C., Roush, T. L., Smith, M. D., 1998. Results from the Mars Global Surveyor Thermal Emission Spectrometer. Science 279, 1692.

Christensen, P. R., Anderson, D. L., Chase, S. C., Clark, R. N., Kieffer, H. H., Malin, M. C., Pearl, J. C., Carpenter, J., Bandiera, N., Brown, F. G., 1992. Thermal emission spectrometer experiment - Mars Observer mission. J. Geophys. Res. 97, 7719–7734.

Christensen, P. R., Bandfield, J. L., Bell, J. F., III, Gorelick, N., Hamilton, V. E., Ivanov, A., Jakosky, B. M., Kieffer, H. H., Lane, M. D., Malin, M. C., McConnochie, T., McEwen, A. S., McSween, H. Y., Mehall, G. L., Moersch, J. E., Nealson, K. H., Rice, J. W., Richardson, M. I., Ruff, S. W., Smith, M. D., Titus, T. N., Wyatt, M. B., 2003. Morphology and Composition of the Surface of Mars: Mars Odyssey THEMIS Results. Science 300, 2056–2061.

Christensen, P. R., Jakosky, B. M., Kieffer, H. H., Malin, M. C., McSween, H. Y., Jr., Nealson, K., Mehall, G. L., Silverman, S. H., Ferry, S., Caplinger, M., Ravine, M., 2004. The Thermal Emission Imaging System (THEMIS) for the Mars 2001 Odyssey Mission. Space Science Reviews 110, 85–130.

Christensen, P. R., McSween, H. Y., Bandfield, J. L., Ruff, S. W., Rogers, A. D., Hamilton, V. E., Gorelick, N., Wyatt, M. B., Jakosky, B. M., Kieffer, H. H., Malin, M. C., Moersch, J. E., 2005. Evidence for magmatic evolution and diversity on Mars from infrared observations. Nature 436, 504–509.

Clancy, R. T., Sandor, B. J., 1998. CO_2 ice clouds in the upper atmosphere of Mars. Geophys. Res. Lett. 25, 489–+.

Clancy, R. T., Sandor, B. J., Wolff, M. J., Christensen, P. R., Smith, M. D., Pearl, J. C., Conrath, B. J., Wilson, R. J., 2000. An intercomparison of ground-based millimeter, MGS TES, and Viking atmospheric temperature measurements: Seasonal and interannual variability of temperatures and dust loading in the global Mars atmosphere. J. Geophys. Res. 105 (14), 9553–9572.

Clancy, R. T., Wolff, M. J., Christensen, P. R., 2003. Mars aerosol studies with the MGS TES emission phase function observations: Optical depths, particle sizes,

and ice cloud types versus latitude and solar longitude. Journal of Geophysical Research (Planets) 180, 2–1.

Clancy, R. T., Wolff, M. J., Whitney, B. A., Cantor, B. A., 2006. Mars equatorial mesospheric clouds. In: Forget, F., Lopez-Valverde, M. A., Desjean, M. C., Huot, J. P., Lefevre, F., Lebonnois, S., Lewis, S. R., Millour, E., Read, P. L., Wilson, R. J. (Eds.), Second workshop on Mars atmosphere modelling and observations,, pp. 264–+.

Colaprete, A., Haberle, R. M., Toon, O. B., 2003. Formation of convective carbon dioxide clouds near the south pole of Mars. Journal of Geophysical Research (Planets) 108, 17.

Colaprete, A., Toon, O. B., 2002. Carbon dioxide snow storms during the polar night on Mars. J. Geophys. Res. 107, 5.

Colaprete, A., Toon, O. B., 2003. Carbon dioxide clouds in an early dense Martian atmosphere. Journal of Geophysical Research (Planets) 108, 6–1.

Conrath, B. J., Pearl, J. C., Smith, M. D., Maguire, W. C., Christensen, P. R., Dason, S., Kaelberer, M. S., 2000. Mars Global Surveyor Thermal Emission Spectrometer (TES) observations: Atmospheric temperatures during aerobraking and science phasing. J. Geophys. Res. 105, 9509–9520.

Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Encrenaz, T., 2001. The Atmosphere of Mars as Constrained by Remote Sensing. Space Science Reviews 96, 411–424.

Gaddis, L., Anderson, J., Becker, K., Becker, T., Cook, D., Edwards, K., Eliason, E., Hare, T., Kieffer, H., Lee, E. M., Mathews, J., Soderblom, L., Sucharski, T., Torson, J., McEwen, A., Robinson, M., 1997. An Overview of the Integrated Software for Imaging Spectrometers (ISIS). In: Lunar and Planetary Institute Conference Abstracts, pp. 387–+.

Glandorf, D. L., Colaprete, A., Tolbert, M. A., Toon, O. B., 2002. CO_2 Snow on Mars and Early Earth: Experimental Constraints. Icarus 160, 66–72.

Haberle, R. M., Pollack, J. B., Barnes, J. R., Zurek, R. W., Leovy, C. B., Murphy, J. R., Lee, H., Schaeffer, J., 1993. Mars atmospheric dynamics as simulated by the NASA AMES General Circulation Model. I - The zonal-mean circulation. J. Geophys. Res. 98, 3093–3123.

Hansen, G. B., 1997. Spectral absorption of solid CO_2 from the ultraviolet to the far-infrared. Advances in Space Research 20, 1613–1616.

Hartmann, D. L., 1983. Barotropic Instability of the Polar Night Jet Stream. Journal of Atmospheric Sciences 40, 817–835.

Henyey, L. G., Greenstein, J. L., 1941. Diffuse radiation in the Galaxy. Astrophys. J. 93, 70–83.

Hinson, D. P., Simpson, R. A., Twicken, J. D., Tyler, G. L., Flasar, F. M., 1999. Initial results from radio occultation measurements with Mars Global Surveyor. J. Geophys. Res. 104 (13), 26997.

Holton, J. R., 1983. The Influence of Gravity Wave Breaking on the General Circulation of the Middle Atmosphere. Journal of Atmospheric Sciences 40, 2497–2507.

Hunten, D. M., Turco, R. P., Toon, O. B., 1980. Smoke and dust particles of meteoric origin in the mesosphere and stratosphere. Journal of Atmospheric Sciences 37, 1342–1357.

Ivanov, A. B., Muhleman, D. O., 2001. Cloud Reflection Observations: Results from the Mars Orbiter Laser Altimeter. Icarus 154, 190–206.

Jakosky, B. M., Haberle, R. M., 1992. The seasonal behavior of water on Mars. In: Mars. The University of Arizona Press, Tucson, Arizona, pp. 969–1016.

Jakosky, B. M., Martin, T. Z., 1987. Mars - North-Polar atmospheric warming during dust storms. Icarus 72, 528–534.

James, P. B., Kieffer, H. H., Paige, D. A., 1992. The seasonal cycle of carbon dioxide on Mars. In: Kieffer, H. H., Jakosky, B. M., Snyder, C. W., Matthews, M. S. (Eds.), Mars. The University of Arizona Press, Tucson, Arizona, pp. 934–968.

Jaquin, F., 1988. Very high elevation water ice clouds on Mars: Their morphology and temporal behavior. In: Lunar and Planetary Inst., MECA Workshop on Atmospheric H2O Observations of Earth and Mars. Physical Processes, Measurements and Interpretations p 58-61 (SEE N89-25790 19-91), pp. 58–61.

Jaquin, F., Gierasch, P., Kahn, R., 1986. The vertical structure of limb hazes in the Martian atmosphere. Icarus 68, 442–461.

Jaquin, R. F. I., 1989. *The middle Martian atmosphere*. Ph. D. thesis, Cornell University.

Joshi, M. M., Lawrence, B. N., Lewis, S. R., 1995. Gravity wave drag in three-dimensional atmospheric models of Mars. J. Geophys. Res. 100, 21235–21246.

Kahn, R. A., Martin, T. Z., Zurek, R. W., Lee, S. W., 1992. *The Martian dust cycle*, pp. 1017–1053. The University of Arizona Press, Tucson, Arizona.

Klostermeyer, J., 2002. Noctilucent clouds getting brighter. Journal of Geophysical Research (Atmospheres) 107, 1–1.

Kokhanovsky, A. A., 2005. Microphysical and optical properties of noctilucent clouds. Earth Science Reviews 71, 127–146.

Lait, L. R., 1994. An Alternative Form for Potential Vorticity. Journal of Atmospheric Sciences 51, 1754–1759.

Lellouch, E., Rosenqvist, J., Goldstein, J. J., Bougher, S. W., Paubert, G., 1991. First absolute wind measurements in the middle atmosphere of Mars. Astrophysical Journal 383, 401–406.

Magalhães, J. A., Schofield, J. T., Seiff, A., 1999. Results of the Mars Pathfinder atmospheric structure investigation. J. Geophys. Res. 104 (13), 8943–8956.

Malin, M. C., Calvin, W., Clancy, R. T., Haberle, R. M., James, P. B., Lee, S. W., Thomas, P. C., Caplinger, M. A., 2001. The Mars Color Imager (MARCI) on the Mars Climate Orbiter. J. Geophys. Res. 106, 17651–17672.

Malin, M. C., Danielson, G. E., Ingersoll, A. P., Masursky, H., Veverka, J., Ravine, M. A., Soulanille, T. A., 1992. Mars Observer Camera. J. Geophys. Res. 97, 7699–7718.

Martin, T. Z., Kieffer, H. H., 1979. Thermal infrared properties of the Martian atmosphere. II - The 15-micron band measurements. J. Geophys. Res. 84, 2843–2852.

McIntyre, M. E., Palmer, T. N., 1983. Breaking planetary waves in the stratosphere. Nature 305, 593–600.

Milam, K. A., Stockstill, J. E., Moersch, J. E., McSween, H. Y., Tornabene, L. L., Ghosh, A., Wyatt, M. B., Christensen, P. R., 2003. THEMIS characterization of the MER Gusev crater landing site. J. Geophys. Res. 108, 8078.

Nash, E. R., Newman, P. A., Rosenfield, J. E., Schoeberl, M. R., 1996. An objective determination of the polar vortex using Ertel's potential vorticity. J. Geophys. Res. 101, 9471–9478.

Neukum, G., Jaumann, R., 2004. HRSC: The High Resolution Stereo Camera of Mars Express. In: Wilson, A., Chicarro, A. (Eds.), Mars Express: The Scientific Payload, Number 1240 in SP, Noordwijk, Netherlands, pp. 17–35. ESA, ESA Publications Division.

Orton, G. S., Fisher, B. M., Baines, K. H., Momary, T. W., Fox, O. D., 2002. A Cold Hole at the Pole of Jupiter. Bulletin of the American Astronomical Society 34, 855.

Owen, T., 1992. The composition and early history of the atmosphere of Mars. In: Kieffer, H. H., Jakosky, B. M., Snyder, C. W., Matthews, M. S. (Eds.), Mars. The University of Arizona Press, Tucson, Arizona, pp. 818–834.

Pearl, J. C., Smith, M. D., Conrath, B. J., Bandfield, J. L., Christensen, P. R., 2001. Observations of Martian ice clouds by the Mars Global Surveyor Thermal Emission Spectrometer: The first Martian year. J. Geophys. Res. 106, 12325–12338.

Pelkey, S. M., Jakosky, B. M., Christensen, P. R., 2003. Surficial properties in Melas Chasma, Mars, from Mars Odyssey THEMIS data. Icarus 165, 68–89.

Pelkey, S. M., Jakosky, B. M., Christensen, P. R., 2004. Surficial properties in Gale Crater, Mars, from Mars Odyssey THEMIS data. Icarus 167, 244–270.

Pollack, J. B., Haberle, R. M., Schaeffer, J., Lee, H., 1990. Simulations of the general circulation of the Martian atmosphere. I - Polar processes. J. Geophys. Res. 95, 1447–1473.

Pollack, J. B., Leovy, C. B., Greiman, P. W., Mintz, Y., 1981. A Martian general circulation experiment with large topography. Journal of Atmospheric Sciences 38, 3–29.

Polvani, L. M., Saravanan, R., 2000. The Three-Dimensional Structure of Breaking Rossby Waves in the Polar Wintertime Stratosphere. Journal of Atmospheric Sciences 57, 3663–3685.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992. *Numerical recipes in C. The art of scientific computing*. Cambridge University Press, 2nd ed.

Randel, W. J., 1987. The Evaluation of Winds from Geopotential Height Data in the Stratosphere. Journal of Atmospheric Sciences 44, 3097–3120.

Richardson, M. I., Inada, A., Strausberg, M. J., McConnochie, T. H., Wang, H., 2006. High-resolution atmospheric observations by the Mars Odyssey Thermal Emission Imaging System. Revised and resubmitted to J. Geophys. Res.

Schoeberl, M. R., Hartmann, D. L., 1991. The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions. Science 251, 46–52.

Schoeberl, M. R., Lait, L. R., Newman, P. A., Rosenfield, J. E., 1992. The Structure of the Polar Vortex. J. Geophys. Res. 97, 7859–7882.

Smith, M. D., 2004. Interannual variability in TES atmospheric observations of Mars during 1999-2003. Icarus 167, 148–165.

Smith, M. D., Bandfield, J. L., Christensen, P. R., Richardson, M. I., 2003. Thermal Emission Imaging System (THEMIS) infrared observations of atmospheric dust and water ice cloud optical depth. Journal of Geophysical Research (Planets) 108, 1–1.

Smith, M. D., Pearl, J. C., Conrath, B. J., Christensen, P. R., 2000. Mars Global Surveyor Thermal Emission Spectrometer (TES) observations of dust opacity during aerobraking and science phasing. J. Geophys. Res. 105, 9539–9552.

Smith, M. D., Pearl, J. C., Conrath, B. J., Christensen, P. R., 2001. Thermal Emission Spectrometer results: Mars atmospheric thermal structure and aerosol distribution. J. Geophys. Res. 106, 23929–23946.

Smith, P. H., Bell, J. F., III, Bridges, N. T., Britt, D. T., Gaddis, L., Greeley, R., Keller, H. U., Herkenhoff, K. E., Jaumann, R., Johnson, J. R., Kirk, R. L., Lemmon, M., Maki, J. N., Malin, M. C., Murchie, S. L., Oberst, J., Parker, T. J., Reid, R. J., Sablotny, R., Soderblom, L. A., Stoker, C., Sullivan, R., Thomas, N., Tomasko, M. G., Ward, W., Wegryn, E., 1997. Results from the Mars Pathfinder Camera. Science 278, 1758–+.

Soderblom, J. M., Bell, J. F., III, Hubbard, M. Y. H., Wolff, M. J., 2005. Martian phase function: Modeling the visible to near-infrared surface photometric function using HST-WFPC2 data. submitted to Icarus.

Sonnabend, G., Wirtz, D., Vetterle, V., Schieder, R., 2005. High-resolution observations of Martian non-thermal CO_2 emission near 10 μm with a new tuneable heterodyne receiver. Astronomy and Astrophysics 435, 1181–1184.

Squyres, S. W., Arvidson, R. E., Bell, J. F., III, Brückner, J., Cabrol, N. A., Calvin, W., Carr, M. H., Christensen, P. R., Clark, B. C., Crumpler, L., Des Marais, D. J., d'Uston, C., Economou, T., Farmer, J., Farrand, W., Folkner, W., Golombek, M., Gorevan, S., Grant, J. A., Greeley, R., Grotzinger, J., Haskin, L., Herkenhoff, K. E., Hviid, S., Johnson, J., Klingelhöfer, G., Knoll, A., Landis, G., Lemmon, M., Li, R., Madsen, M. B., Malin, M. C., McLennan, S. M., McSween, H. Y., Ming, D. W., Moersch, J., Morris, R. V., Parker, T., Rice, J. W., Richter, L., Rieder, R., Sims, M., Smith, M., Smith, P., Soderblom, L. A., Sullivan, R., Wänke, H., Wdowiak, T., Wolff, M., Yen, A., 2004a. The Spirit Rover's Athena Science Investigation at Gusev Crater, Mars. Science 305, 794–800.

Squyres, S. W., Arvidson, R. E., Bell, J. F., III, Brückner, J., Cabrol, N. A., Calvin, W., Carr, M. H., Christensen, P. R., Clark, B. C., Crumpler, L., Des Marais, D. J., d'Uston, C., Economou, T., Farmer, J., Farrand, W., Folkner, W., Golombek, M., Gorevan, S., Grant, J. A., Greeley, R., Grotzinger, J., Haskin, L., Herkenhoff, K. E., Hviid, S., Johnson, J., Klingelhöfer, G., Knoll, A. H., Landis, G., Lemmon, M., Li, R., Madsen, M. B., Malin, M. C., McLennan, S. M., McSween, H. Y., Ming, D. W., Moersch, J., Morris, R. V., Parker, T.,

Rice, J. W., Richter, L., Rieder, R., Sims, M., Smith, M., Smith, P., Soderblom, L. A., Sullivan, R., Wänke, H., Wdowiak, T., Wolff, M., Yen, A., 2004b. The Opportunity Rover's Athena Science Investigation at Meridiani Planum, Mars. Science 306, 1698–1703.

Stamnes, K., Tsay, S.-C., Jayaweera, K., Wiscombe, W., 1988. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. Appl. Opt. 27, 2502–2509.

Taylor, F. W., 2002. Some fundamental questions concerning the circulation of the atmosphere of Venus. Advances in Space Research 29, 227–231.

Thompson, D. W. J., Wallace, J. M., 2000. Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability*. Journal of Climate 13, 1000–1016.

Thompson, D. W. J., Wallace, J. M., Hegerl, G. C., 2000. Annular Modes in the Extratropical Circulation. Part II: Trends. Journal of Climate 13, 1018–1036.

Titus, T. N., Kieffer, H. H., Christensen, P. R., 2003. Exposed Water Ice Discovered near the South Pole of Mars. Science 299, 1048–1051.

Toigo, A., Richardson, M. I., Newman, C. E., 2005. Planetary WRF: a Multi-Scale, Planetary, Atmospheric Model. In: AGU Fall Meeting Abstracts, pp. A140+.

Torson, J. M., Becker, K. J., 1997. ISIS - A Software Architecture for Processing Planetary Images. In: Lunar and Planetary Institute Conference Abstracts, pp. 1443–+.

Wang, H., Ingersoll, A. P., 2002. Martian clouds observed by Mars Global Surveyor Mars Orbiter Camera. Journal of Geophysical Research (Planets) 110 (8), 8–1.

Wang, H., Zurek, R. W., Richardson, M. I., 2005. Relationship between frontal dust storms and transient eddy activity in the northern hemisphere of Mars as observed by Mars Global Surveyor. Journal of Geophysical Research (Planets) 110 (E9), 7005–+.

Warren, S. G., 1984. Optical constants of ice from the ultraviolet to the microwave. Applied Optics 23, 1206–1225.

Waugh, D. W., Randel, W. J., 1999. Climatology of Arctic and Antarctic Polar Vortices Using Elliptical Diagnostics. Journal of Atmospheric Sciences 56, 1594–1613.

Wilson, R. J., 1997. A general circulation model simulation of the Martian polar warming. Geophys. Res. Lett. 24, 123.

Wilson, R. J., Banfield, D., Conrath, B. J., Smith, M. D., 2002. Traveling waves in the Northern Hemisphere of Mars. Geophys. Res. Lett. 29, 29.

Wolff, M. J., 11 co-authors, 2006. Constraints on Dust Aerosols from the Mars Exploration Rovers Using MGS Overflights and Mini-TES. submitted to J. Geophys. Res.

Wolff, M. J., James, P. B., Todd Clancy, R., Lee, S. W., 1999. Hubble Space Telescope observations of the Martian aphelion cloud belt prior to the Pathfinder mission: Seasonal and interannual variations. J. Geophys. Res. 104, 9027–9042.

Zurek, R. W., 1992. Comparative aspects of the climate of Mars - an introduction to the current atmosphere. In: Mars. The University of Arizona Press, Tucson, Arizona, pp. 799–817.

Zurek, R. W., Barnes, J. R., Haberle, R. M., Pollack, J. B., Tillman, J. E., Leovy, C. B., 1992. Dynamics of the Atmosphere of Mars. In: Kieffer, H. H., Jakosky, B. M., Snyder, C. W., Matthews, M. S. (Eds.), Mars. The University of Arizona Press, Tucson, Arizona, pp. 835–933.