ELSEVIER

Journal of Quantitative Spectroscopy & Radiative Transfer 85 (2004) 367–383

Journal of Quantitative Spectroscopy & Radiative Transfer

www.elsevier.com/locate/jqsrt

Absolute, spectrally-resolved, thermal radiance: a benchmark for climate monitoring from space

J.G. Anderson[a,*], J.A. Dykema[a], R.M. Goody[a], H. Hu[a], D.B. Kirk-Davidoff[b]

[a] *Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA*

[b] *Department of Meteorology, University of Maryland, College Park, MD 20742, USA*

Received 21 February 2003; accepted 23 May 2003

Abstract

Spectrally resolved thermal radiances measured from orbit with an accuracy in brightness temperature of 100 mK constitute a critical observation for climate monitoring. The design of a small, low-cost instrument capable of accuracies of better than 100 mK, demonstrated on-orbit, is presented and analyzed. It is shown that systematic and random errors inherent in observations from space can be reduced to levels commensurate with the instrumental accuracy of 100 mK. Monitoring spectrally resolved radiance, accurate to 100 mK, is feasible, and constitutes a versatile climate Benchmark observation that is needed in the national research strategy.

Keywords: Climate monitoring; Infrared spectroscopy; Satellite observations; Sampling errors

1. Climate monitoring requirements

Spectrally resolved measurements of the outgoing thermal radiation to space have, for the past 30 years, occupied an important place in meteorological observing. In addition to a diagnostic function for weather, this radiation stream is also the output from the climate heat engine and is demonstrably as important for climate as for weather. However, requirements for climate and weather observing differ in important ways. The purpose of this paper is to show how climate requirements can be

* Corresponding author. Tel.: +1-617-495-5922; fax: +1-617-495-4902.
E-mail address: anderson@huarp.harvard.edu (J.G. Anderson).

0022-4073/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0022-4073(03)00232-2

satisfied with simple, state-of-the-art instrumentation if attention is given to these requirements. The alternative of modifying meteorological requirements to include climate requirements is likely to be both expensive and unsatisfactory.

From the standpoint of climate modeling the requirement is for accurate and reliable averages over significant regions of the Earth's surface. These averages must be of sufficient accuracy to resolve questions that arise in climate modeling and, because climate prediction will be a continuing engagement, this accuracy must be convincingly demonstrable to future generations of climate scientists.

Physical process studies and weather forecasting both require simultaneous, bore-sighted observations by several different instruments, for which sensitivity is usually more important than accuracy and long-term stability. An emphasis on climate averages removes the requirement for simultaneous bore-sighted observations, but it introduces the problem of how accurate averages can be obtained from space (the sampling problem), given sufficiently accurate measurements in the first place.

In this paper we shall consider both the instrumental requirements that define the optical design and the sampling requirements in order to present a complete picture of an observing system designed to meet climate requirements. We consider thermal radiance only. Similar considerations apply to solar radiances, direct and scattered, and to microwave radiances.

With respect to the accuracy required for measurements of thermal radiation, the question cannot be fully answered until the use of the data for climate modeling is fully understood. This can only be done in the context of an operational climate model, something that does not exist in the U.S. at the present time [1]. However, the fact that the climate debate involves questions of tenth degree changes [2], suggests that one tenth of a degree (100 mK) is an important and useful level of accuracy. In terms of the integrated thermal radiation from a 250 K blackbody (corresponding approximately to the mean brightness temperature of Earth), one tenth of a degree represents 0.27 W m^{-2}. In the climate debate, 1 W m^{-2} is considered to be important.

The spectral resolution adopted by modern, space-based, meteorological sounders is $\sim$1 cm^{-1}. This spectral resolution is a practical compromise for meteorological data: high enough to yield most information that is likely to be useful for weather forecasting; but the next step is a very large one, to fully resolved line profiles, requiring a resolution of $\sim$0.03 cm^{-1}. Again, a final judgment can only be made in the context of an operational climate model, but the ground is well enough explored by meteorologists to adopt 1 cm^{-1} as a minimum, but useful resolution requirement. Resolution on this order is also required to perform accurate calibrations [3], and for optical system diagnosis.

The required resolution and the required accuracy are relatively straightforward to realize in the laboratory and the results can be referred to international standards in a convincing way. The experimental objective in space is to achieve the same result in a satellite that is unattended for five or more years and, if it is to be the basis of a climate monitoring system, to do so at reasonable cost.

In this paper we bring together all aspects of a space climate Benchmark system based on spectrally-resolved, thermal radiances. For details, three specialized papers are in preparation: Dykema et al. [4] deals with the instrument and its performance; Kirk-Davidoff et al. [5] deals with aliasing the diurnal signal; Hu et al. [6] consider other aspects of sampling from orbit.

2. An FTS for climate monitoring

2.1. Instrument configuration

The crucial requirement for climate monitoring from space is to achieve the required accuracy and, even more importantly, to demonstrate in flight that this accuracy has been achieved. This section describes a system based upon simple Fourier transform spectrometers (FTSs) that satisfies this requirement. It is based upon three principles:

- simplicity of design: the fundamental comparison between black-body standards and the observed scene should be as uncomplicated as possible,
- redundancy in essential systems, combined with on-board diagnostics and
- parallel laboratory and field activities together with continuous engagement of the climate community in a critique of procedures and results.

The instrument configuration discussed here employs two small FTSs. The two instruments are bore-sighted on the Earth's surface. Each has its own electronics, can view space over a 45° range of angles and has two independent black bodies. Several in-flight diagnostic features are included. The key attributes of the instruments are shown in Table 1.

The experimental physics and technology that underlie this design have been developed by many researchers. The key elements include research on thermodynamic [8] and practical thermometry [9], blackbody standards [10–12], calibration procedures [13–16], and detector developments that improve linearity [17,18]. In addition, the special requirements of climate monitoring admit simplified optical designs by minimizing the number of optical elements, diminishing alignment sensitivity, compensating for thermal gradients, and reducing off-axis response. Simplicity of design facilitates the inclusion of redundant calibrations, a pragmatic means to achieve climate Benchmark requirements [19].

The particular instrument configuration discussed in this paper (Fig. 1) draws upon studies of small satellites [20], atmospheric spectroscopy [21], redundant calibrations [22], flight intercomparisons [23] and detector technologies [24].

2.2. Blackbodies

During normal operation, the nadir-viewed spectra will be calibrated from the deep space view and a "hot" blackbody by the method of Revercomb et al. [25]. The deep space view provides an

Table 1
Key instrument attributes

Attribute	Value
FTSs	Two
Instrument absolute accuracy	< 0.1 K
Spectral range	225–1600 cm^{-1}
Spectral resolution (unapodized)	0.6 cm^{-1}
Aperture	2.5 cm
Field of view (half angle)	65 mrad
Footprint	100 km

Fig. 1. Instrument configuration. A key design feature is the use of dual FTSs bore-sighted on the same nadir footprint. Each instrument is a four-port, cube-corner FTS, each with two blackbodies that overdetermine the calibration, and with a deep space view over 45° to calibrate polarization errors.

accurate characterization of the instrument's zero offset, ensuring good calibration for low brightness regions such as the center of the 15 μm CO_2 band. The second blackbody is used to check the radiometric performance.

Blackbody accuracy depends on the surface emissivity (which may deteriorate), the aspect ratio of the cavity aperture, the evenness of temperature within the cavity, and the accuracy of temperature measurement. All of these factors have been treated in detail by Mason et al. [10] in conjunction with the design of the along track scanning radiometer (ATSR) of the European Space Agency. Their blackbody had a design error of < 35 mK from the emissivity, with measured errors of 20 mK from temperature inhomogeneity, and < 10 mK from temperature readout, for a total of < 41 (< 65) mK (the first figure is the square root of the summed squares and the figure in parentheses is a linear sum; the error is expected to lie between these two limits). Blackbodies following this design have been used in space since 1991 and reports on the overall performance of ATSR [26] suggest that the on-orbit performance is almost within a factor of two of the laboratory performance.

The ATSR blackbody design can be improved [27] when coupled to an optical instrument design selected for accuracy. A smaller aperture, with an aspect ratio of 6, reduces the emissivity error to 9 mK and improves temperature homogeneity. These two error sources together contribute < 20 (< 29) mK to the error budget of an improved design.

2.3. The dispersive system

An FTS was chosen as the dispersive system because it has a simple design, is mechanically robust, has a small number of optical components and, of critical importance, has the same instrumental profile at all frequencies. The resolution of $0.6\,\mathrm{cm}^{-1}$ is sufficiently high that line positions and profiles can be checked in flight from atmospheric features. A single mirror external to the FTS selects between the blackbodies, deep space, and a nadir view of the Earth's surface.

Two possible sources of error are: stray radiation; and polarization errors from the scene selection mirror. Stray radiation, as discussed in Jarecke et al. [28], must be explicitly considered for accuracy at the 100 mK level, as it can introduce errors that are significant in a high-accuracy context if it changes during an observation/calibration cycle. Design precautions will be taken both to minimize and to measure this component of the error budget. This error can be reduced by thermal control of the aperture stop, and it can be monitored during flight.

The optical properties of the pointing mirror and the beamsplitter are polarization-dependent. The rotation of the pointing mirror changes the angle between the plane of incidence of the pointing mirror and the plane of incidence at the beamsplitter, introducing a scan angle-dependent, radiometric offset proportional to the contrast in equivalent temperature between the input scene and the mirror. This polarization effect has been discussed for AIRS (the Atmospheric InfraRed Spectrometer) by Aumann and Overoye [29] and for CrIS (Crosstrack Infrared Spectrometer) by Stumpf and Overbeck [30]. The constant of proportionality is, apart from a normalization factor, equal to the product of the polarization coefficients of the mirror and the beamsplitter. Uncorrected, this offset can lead to radiometric errors as large as 2 K, but a simple correction based on optical properties measured on-orbit can be made [4]. Environmental factors, including outgassing and atomic oxygen exposure, may cause these coefficients of polarization to change over time. To maintain the radiometric performance, the measurement of deep space over a 45° range of angles permits an independent determination of the correction parameters throughout the lifetime of the mission, limiting the radiometric error to 33 mK or less for the mission lifetime.

2.4. Detector chain nonlinearities

Nonlinearities can occur in both the detector [31] and in the electronics [29]. AIRS employs a detector array in which each detector receives only a small radiation flux, and detector nonlinearities are negligible. However, the integrated, miniaturized, detector electronics on AIRS can lead to uncorrected nonlinearity errors as high as 1700 mK [29]. On-orbit measurements allow corrections to be applied.

For an FTS the appropriate detector technologies do not require miniaturized electronics and very small nonlinearities are possible. Nonlinearity can be measured in flight and a correction applied (but this could still be the largest source of instrument error). However, in the choice of a detector, it is possible to take advantage of the averaging that must take place for climate purposes (see Section 3). Cooled semiconductor detectors are not required and less sensitive, intrinsically linear, pyroelectric detectors can be used. We are left with the nonlinearity of precision electronics, $< 0.01\%$. An additional advantage of a pyroelectric detector is sensitivity to radiation at $225\,\mathrm{cm}^{-1}$, which includes much of the water vapor rotation band (important because it gives information on the middle-to-upper

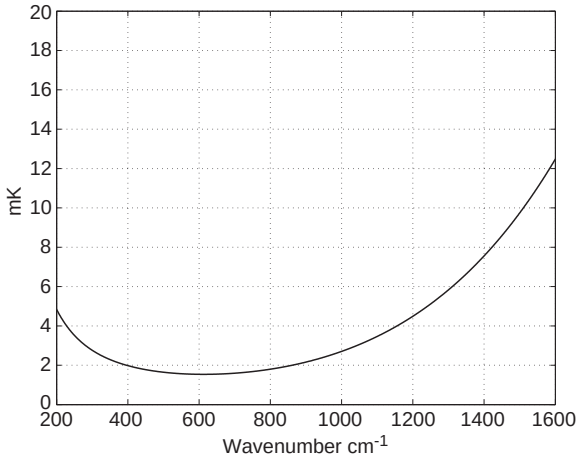

Fig. 2. Errors associated with a nonlinearity of 0.01% for a 250 K blackbody.

Table 2
Sources of systematic error in the FTS system

Subsystem	Source of largest error	Magnitude (mK)
Blackbodies	Thermometry	<24 (<39)
FTS	Polarization	33
Detector chain	Nonlinearities	<14
Other errors		<10
TOTAL		<44.5 (<96)

Sources of systematic error in the FTS system for observations of a 250 K blackbody at 900 cm^{-1}. The first figure in the third column is the square root of the sum of squares. Linear sums of errors are given in parentheses.

troposphere). Errors for a 250 K blackbody and a nonlinearity of 0.01% are shown in Fig. 2, and a budget for all errors is given in Table 2.

Random variations of radiance caused by variable cloud fields are discussed in the following sections. The standard deviation of a single spectrum can be as high as 10–20 K in radiance temperature, and enough observations must be taken to average this variance out. By comparison, the expected noise equivalent radiance from the linear pyroelectric detectors is 0.7 K, and can be neglected.

2.5. System verification

Future weather sounders such as AIRS and CrIS may achieve the accuracies required for climate systems before launch, but there is no assurance of continuation of pre-flight accuracy. The instrument discussed here, on the other hand, places strong emphasis on verification. Apart from a simple and robust design, the present instrument establishes verification through redundancies and in-flight, subsystem monitoring. Such high levels of precaution will provide a standard to establish Benchmark

data in perpetuity. The important redundancies are:

- The highest level of redundancy is the use of the two instruments bore-sighted on the Earth's surface. The most discriminating demonstration of the consistency of this observing system is that these two independent instruments should agree precisely.
- Each instrument has two identical blackbodies for calibration which cover wide, interchangeable temperature ranges. The linearity of the detectors means that two, independent calibrations are available for each of two, independent FTSs.
- Temperature homogeneity in the blackbodies is monitored by four platinum resistance, and four thermistor thermometers. The design of the blackbodies should lead to a homogenous field that requires only one or two thermometers to establish the radiance.

Critical subsystems can be monitored in flight:

- The blackbody emissivity can easily be monitored by illumination with an out-of-aperture source. The depth of the cavity ensures that the surface reflectivity must more than double to impact the total radiometric error.
- Any on-orbit thermometry drift will be revealed by a mix of negative temperature coefficient (NTC) thermistors and encapsulated platinum resistance thermometers (PRTs), which are based on different electronic conduction processes and employ different topologies of electronic readout.
- A temperature-controlled element behind the entrance aperture provides a strict test for stray radiation entering the FTS off-axis.
- A 45° scan on the space view directly measures changes in the polarization correction.
- End-to-end detector signal chain linearity will be tested at the complementary output ports of the FTS. The modulated parts of the interferogram, recorded simultaneously at the two ports, should sum to zero. Deviation from zero is a measure of nonlinearity in the detector chain, the origin of which can be determined independently on-orbit [32].

3. Sampling errors

3.1. Aliasing the diurnal variation

Sampling the diurnal cycle differs for Sun-synchronous orbits ($\approx 99°$ inclination), polar orbits, and low-inclination, precessing orbits. A polar-orbiting satellite remains in a fixed orbital plane, sampling at two times of day, 12 h apart. Both local times cycle through 24 h in the course of a year as the plane of the satellite orbit rotates with respect to the Earth–Sun line. A Sun-synchronous satellite orbits in a plane that corresponds to the Earth–Sun line, and samples two, fixed, discrete, local times only. A low-latitude orbiter can sample the diurnal cycle many times in a year (six times for an inclination of 33° and an altitude of 662 km), at the expense of limited coverage of the surface.

It is not practical to have all climate satellites in the same Sun-synchronous orbits with the same equator crossing times, nor would this be desirable, because the diurnal variation is not simple and varies over the globe. The only way to ensure Benchmark quality is to adopt an observing strategy that leads to accurate climate means. This gives rise to another problem, namely aliasing the diurnal signal into the climate mean because of the periodicity of satellite measurements.

A theoretical study of aliasing errors for constellations of satellites in Sun-synchronous orbits has been made by Leroy [33] for the case of large-amplitude diurnal variability in surface temperature.

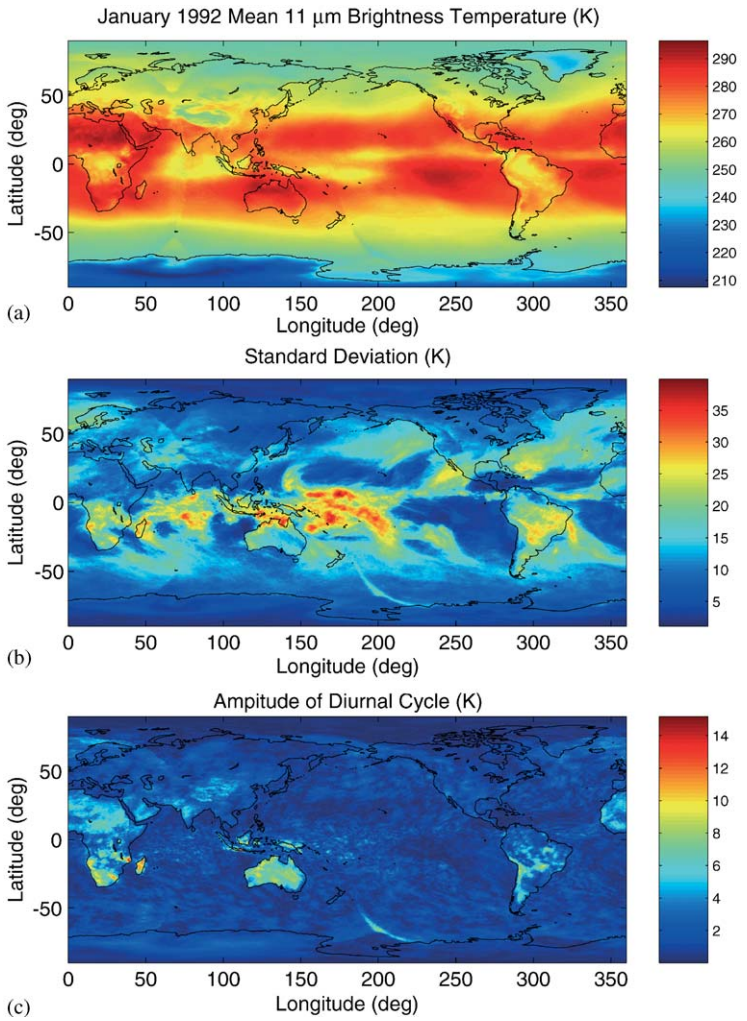

Fig. 3. Data for January 1–16, 1992 from the GCI 11μm data set. (a) Mean brightness temperature. (b) Standard deviation of the brightness temperature. (c) Amplitude of the diurnal cycle.

The following discussion is based on a numerical study of different orbits using the Salby Global Cloud Imagery data set [34]. The GCI Data are regridded 11 μm radiances on a 512×512 grid of pixels that cover 0.35° of latitude and 0.7° of longitude. The variance of 11 μm radiation is close to the maximum in the thermal spectrum (see Fig. 7b), so that calculations using GCI data yield an upper limit to radiance temperature retrieval errors.

A 11μm brightness temperature is a measure of cloud-top temperature, or surface temperature if the skies are clear. As shown in Fig. 3a brightness temperature is a minimum over Antarctica, a local minimum at the equator where deep convection is strongest, and a maximum over the subtropical deserts. Variability (Fig. 3b) is strongest in the equatorial belt, where brightness temperature varies between 190 and 302 K but weakest over the great stratocumulus fields of the subtropical oceans. There are secondary maxima in the mid-latitude storm tracks.

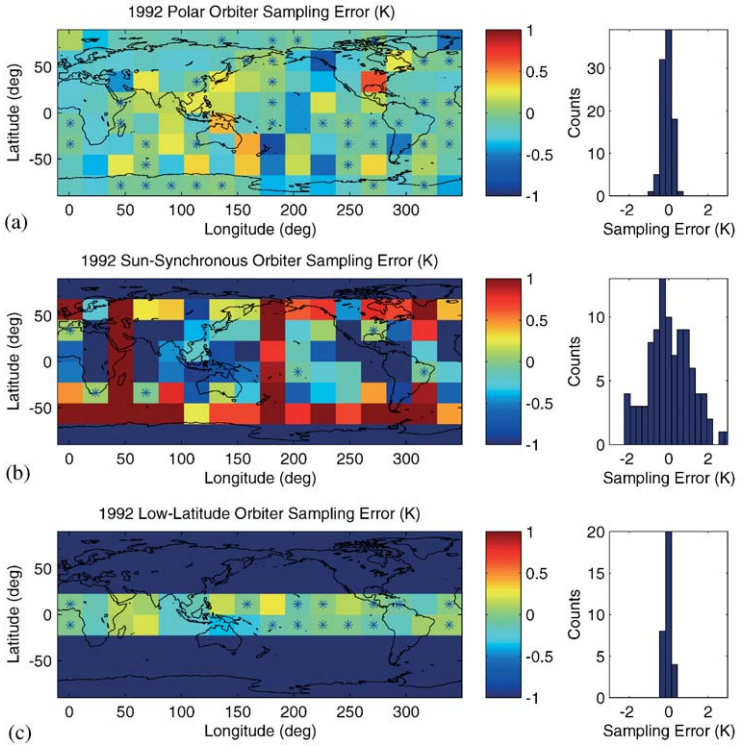

Fig. 4. Retrieval errors for 11μm brightness temperature for single satellites, averaged for 1992. The grid boxes are 22.5° square. A sampling error less than 0.1 K is distinguished with an asterisk (*). Histograms of sampling errors are shown on the right. (a) A single polar orbiter (inclination: 90°, altitude: 833 km). (b) A single Sun-synchronous orbiter (inclination: 98.765°, altitude: 833 km). (c) A single tropical orbiter (inclination: 33°, altitude: 662 km).

Fig. 3c shows the diurnal variation of the brightness temperature. It dominates the total variance for clear skies in desert locations. In other regions, variations in cloud fraction dominate. These random variations are discussed in Section 3.3. Here we are concerned with the systematic errors caused by aliasing.

None of the single satellites in Fig. 4 achieves 100 mK accuracy for all occasions, although the tropical orbiter comes close. Three polar orbiters (Fig. 5b) give still higher accuracy than a single tropical orbiter (Fig. 4c) and at all latitudes and longitudes. Errors are always less than 210 mK with 84% of errors less than 100 mK. For three Sun-synchronous satellites on the other hand, the maximum error is 980 mK and only 50% of the errors are less than 100 mK. These errors can be further reduced by averaging. Fig. 6 shows zonal averages of the results in Figs. 4 and 5. Now retrieval errors are less than 100 mK for even a single polar orbiter, although they are still large (up to 1 K) for a single Sun-synchronous orbiter.

The sparse sampling of the diurnal cycle of brightness temperature from a small number of orbiting radiometers places an important constraint on the accuracy of Benchmark radiance measurements. Accuracy close to 100 mK is only obtainable with three polar orbiters or, for the tropical regions only, with a single low-latitude orbiter. The canonical meteorological constellation of three Sun-synchronous orbiters (Fig. 5c) clearly compromises climate investigations. Averages for

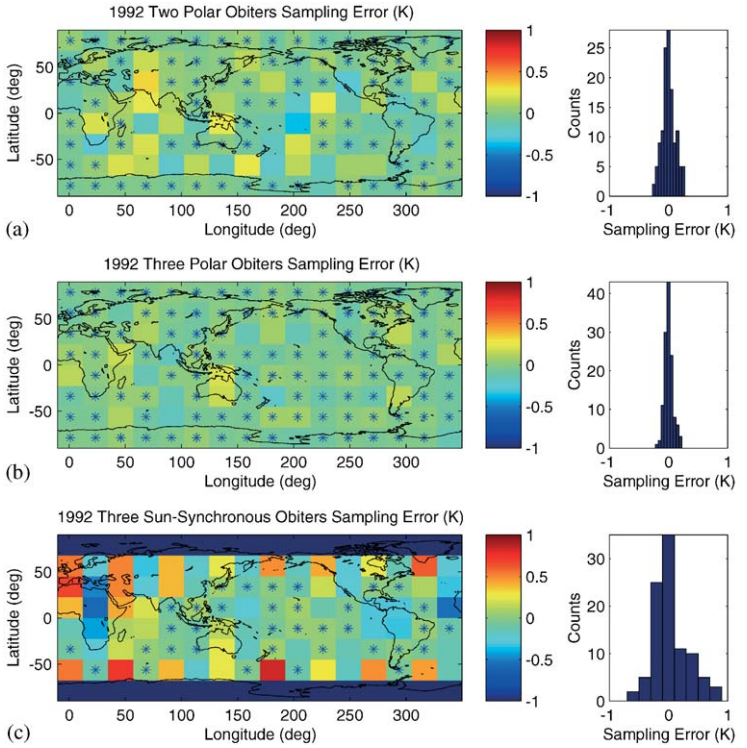

Fig. 5. As for Fig. 4 but for constellations of orbiters, equally spaced in longitude. (a) Two polar orbiters. (b) Three polar orbiters. (c) Three Sun-synchronous orbiters.

individual climate regions can show very large errors that are only partially reproducible from year-to-year. If these reproducible errors are eliminated, or if zonal averages are taken, the meteorological constellation can yield averages as accurate as 200 mK.

These limitations only apply to the large diurnal variation of radiation from clouds or from the Earth's surface. Radiation from an absorption band that originates from above the clouds, or microwave radiances that are not influenced by clouds, will have smaller aliasing errors.

3.2. Footprint errors

IRIS [7] had a large footprint, approximately 95 km. Most missions have smaller footprints, e.g., IMG (the Interferometric Monitor for Greenhouse gases) has a footprint of 8 km [35]. The footprint of a typical GCM, with which data must eventually be compared, is larger (~250 km). But climate data are normally assembled into regions that exceed 1000 km on a side. If the data points are sufficient in number, and evenly distributed over the climate region, the footprint differences alone will not affect the mean value of the radiance (see Haskins et al. [36], Fig. 2 for an example of averaging with two different footprints).

The footprint size does, however, affect the standard deviation of a single observation. Most of the variance shown in Fig. 7b (to be discussed in Section 3.2) is atmospheric, but the difference between the standard deviations of the two instruments is almost entirely attributable to footprint size.

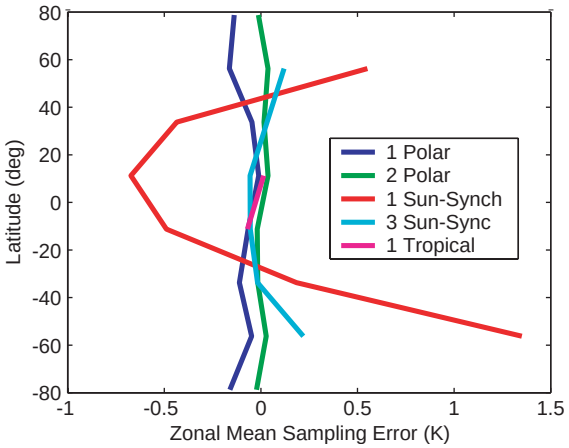

Fig. 6. Zonal averages for orbits from Figs. 4 and 5.

Brindley and Harries [37] evaluated the partial variance for different footprints by an independent method. The partial variance that they obtain is approximately the same as the variance difference between IRIS and IMG in Fig. 7b. Footprint considerations are accounted for if observed standard deviations, and averages over large climate regions, are used.

Brindley and Harries also point to the large footprint errors than can occur when apparently cloud-free regions are selected from the data. This is the reason why it is inadvisable to employ cloud-clearing protocols if precise climate data are the objective.

3.3. Spatial and temporal sampling errors

Harries et al. [38] studied the differences between IRIS and IMG data sets after correcting for the different spectral resolutions (2.8 cm^{-1} for IRIS and 0.10–0.25 cm^{-1} for IMG). Since both missions were very badly sampled in time and space, intercomparison offers a challenging task.

Fig. 7 shows mean radiance spectra and standard deviations for individual IRIS and IMG observations over three comparable time periods totaling 60 days, and covering the entire tropical belt between 30°N and 30°S. The maximum standard deviation for a single spectrum is typically about 12 K for IRIS and 14 K for IMG. This variance is largely due to cloud variations but geographic variations also contribute, as do slow time variations, and unequal numbers of AM and PM spectra.

It is evident from Fig. 7b that to sample the radiation field with an accuracy of 100 to 200 mK requires averages over $\sim 10^4$ independent spectra. A typical climate region in the tropical Pacific Ocean between 10°N and 10°S, and between 130°W and 180°, observed by a polar orbiting satellite that makes an observation every 10 s, will have a total of 6.3×10^3 observed spectra in one month. The samples used in Fig. 7 contained 2.9×10^4 IRIS spectra and 2.6×10^4 IMG spectra. These sample numbers reduce random errors to acceptable levels, but several sources of systematic error remain.

Both IMG and IRIS missions sampled at essentially fixed equator-crossing times. Given equal numbers of spectra at the ascending and descending nodes (AM or PM equator crossing), the diurnal

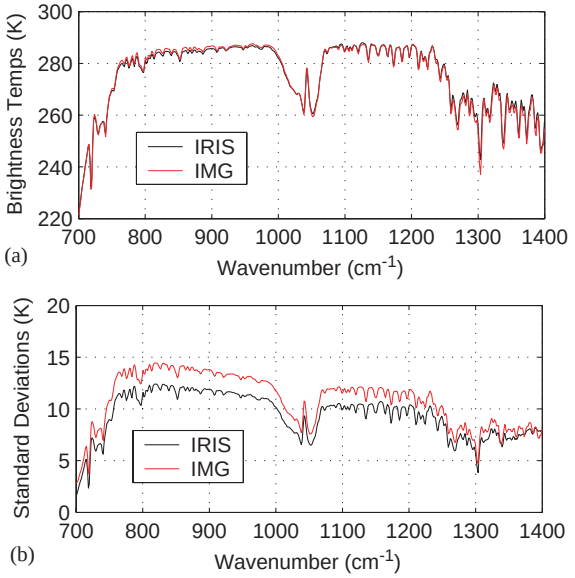

Fig. 7. (a) Means of IRIS and IMG spectra over tropical regions for three 20-day intervals. (b) Standard deviations of single observations. The IMG resolution has been reduced to that of IRIS by modifying the scan length to 35 mm (the IRIS scan), applying the apodization function used by IRIS, and adding a self-apodization function appropriate to the IRIS geometry (Profile P_2 of Steel [39]).

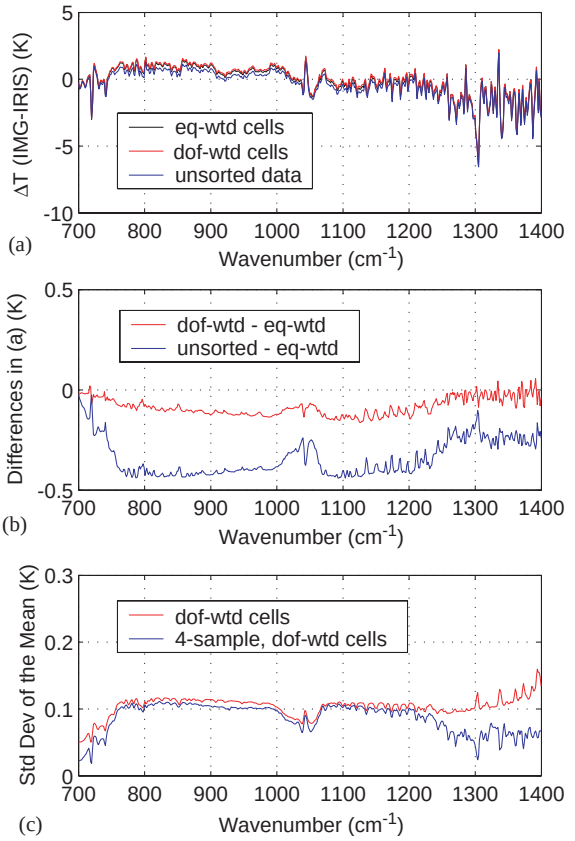

Fig. 8. (a) Three approaches to the average difference $\bar{\Delta}$. (b) shows differences between the curves in a. c. shows two methods of evaluating the standard deviation of the error for dof-weighted cells; 4 equally weighted, random samples, and equal weights for each original observation.

Fourier component is nullified. Because of instrumental problems, this was not the case for either IRIS or IMG, and the diurnal Fourier component can give a systematic error of a few tenths of a degree. Another source of systematic error arises because neither IRIS nor IMG observations were uniformly distributed over the tropics. Since there are geographic differences in radiances, particularly between land and sea, systematic errors can occur. If the two mean spectra in Fig. 7a are subtracted without further precautions, significant systematic errors can result.

To reduce these systematic errors, IRIS and IMG observations were grouped into one of 1296 cells, each of which is expected to be close to homogeneous. Each cell has one of 216, $10^\circ \times 10^\circ$, geographical locations, one of three 20-day periods in which IRIS and IMG observations were both frequent, and either AM or PM equator crossing times. In each cell (subscript i) average IRIS and average IMG radiances were evaluated, and subtracted to form an IMG-IRIS radiance difference Δ_i.

This procedure should lead to no systematic errors within cells or, if small errors remain, they will differ randomly from cell to cell, and should not contribute systematically to the ensemble average.

The average for the entire region, and for the three time periods chosen is

$$\bar{\Delta} = \frac{\sum_i W_i \Delta_i}{\sum_i W_i},$$

where W_i is the weight given to the ith cell. Because Δ_i may vary slightly from cell to cell, this sum may depend on the weights that are used. In Fig. 8a three different possibilities are shown. Equally weighted cells ($W_i = 1$) represents an unbiased estimate of the required average. However, some of the cells contain few observations and are "noisy". This average will have a relatively large standard deviation. *Dof weighted cells* are given a weight proportional to the number of degrees of freedom in the cell. This average will have a smaller standard deviation, but systematic errors are a possibility. The third curve in Fig. 8a (unsorted data) is the difference between the two spectra in Fig. 7a, and is expected to contain biases.

Fig. 8b shows two differences between these three results. Since the equally weighted average is known to have a standard deviation of ~0.25 K, we may conclude that there are significant biases in the unsorted data, but not in the dof-weighted average. We have therefore adopted the dof-weighted case as the best result from these poorly sampled data.

The standard deviation of this mean is as important as the mean itself. In Fig. 8c the standard deviation of the dof-weighted mean is estimated in two different ways. First, by estimating the standard deviation of each cell on the basis that the original spectra are each independent; second, by dividing the original data set into four random subsets and averaging each according to the number of degrees of freedom of each cell.

The results in Fig. 8c show that climate averages can be obtained with useful levels of accuracy, even for very badly sampled data. Useful results are also obtainable with smaller samples, or with smaller geographic regions. If no data are missing, the Central Pacific climate region, discussed above, would give results at least as good as those in Fig. 8c for a seasonal average. The observing system discussed in Section 2, with its two independent instruments will provide uninterrupted data without gaps due to faulty operation.

Brindley and Harries [37,40] have also examined the problem of sampling the differences between IRIS/IMG data. They used much smaller data sets (some with single numbers) and present a picture of large random errors. The present treatment covers all relevant aspects of the Brindley and Harries analysis, and all of the differences between the analyzes are attributable to the different sample sizes.

4. Final comments

The task of the climate analyst only begins when good data become available. The spectra in Fig. 8a are the result of a superposition of two different effects. First, there is an increase of greenhouse gases from 1970 to 1996 that gives rise to recognizable bands in the observed spectrum. This effect is known and understood theoretically and is the basis for the climate forcings of all GCMs. Harries et al. [38] showed that the expected bands could be detected in the observed IRIS/IMG difference spectra, confirming the capabilities of the two observing systems. But the important problem for

modern climate science is to predict and to measure the response of other atmospheric variables (temperature, humidity and cloud) to a climate forcing. These changes also leave characteristic imprints on the outgoing thermal spectrum. The requirement is to separate forcing and response, and to compare the response to theoretical predictions.

The response spectrum emphasizes the window region (800–1200 cm^{-1}). This region is strongly affected by errors in instrument calibration. IRIS and IMG were calibrated before flight, but no details are given, and there is no reason to believe that the calibration was maintained during flight. One degree calibration error shared by the two instruments is not implausible. Consequently, information from Fig. 8a in the window region is of limited value for determining the atmospheric response over 26 years, whereas a system with the measurement accuracy discussed in this paper would yield important information on atmospheric behavior.

The response signal is not caused by forcing alone, but is a combination of forcing with the natural variability of the system. Natural variability in IRIS spectra has been studied by Haskins et al. [41,42]. Techniques of optimal filtering, that can separate the two signals, are known. In addition, natural variability is a property of the climate system that can also be monitored and used to test model performance [43].

We have presented the case for a key climate Benchmark observation, and a technically realistic approach to climate measurement. This advance is based on a remarkable, early instrument development. IRIS was designed in the 1960s and flown in 1970/71. In retrospect, it was far ahead of its time. After IRIS, U.S. satellite radiance measurements evolved along two different paths, one for weather and one for climate.

Weather satellites have employed low-spectral-resolution sounders (microwave and thermal infrared) with intrinsic accuracies of 1 K. These early instruments are about to be replaced by much more complex instruments with higher spectral resolution, but without precautions needed to ensure high accuracy. AIRS, which flies on the NASA mission EOS Aqua, and CrIS, on NPOESS (the National Polar Orbiting Environmental Satellite System), are superbly designed for their purpose. They emphasize small footprints, daily global coverage, short dwell times, and observations at specific local times. None of these factors is required for climate monitoring and all make the fundamental goal of dependable, high accuracy harder to achieve.

Climate research, on the other hand, has been supported by many flights of NASA's Earth Radiation Budget Experiment (ERBE) and its successor, Clouds and the Earth's Radiant Energy System (CERES). These radiometers are designed for accurate measurements of spectrally integrated solar radiation and spectrally integrated thermal radiation. Observations can be calibrated against standard blackbodies only with some prior knowledge of the observed scene [3]. Given sufficient prior information the integrated thermal radiation can be determined to 1%, corresponding to an accuracy for the emission temperature of 600 mK, but this comes at the expense of the statistical independence of radiance and other meteorological parameters.

Even if integrated fluxes could be accurately measured, spectral resolution constitutes an essential requirement for climate monitoring. To understand this, consider a circumstance under which a climate change occurs, but the effects of different spectral regions cancel to give no net change in the integrated spectrum. ERBE or CERES would yield no information, but all of the many changes would be registered in a resolved spectrum.

The climate monitoring system discussed in this paper is in a different cost class from modern meteorological sounders or NASA's climate radiance monitors. A distinguishing feature is the simple,

uncomplicated design, that can lead to a low-cost system. The proposed FTS is a version of IRIS with modern technology. A configuration of this instrument that was proposed to NASA in 1995 weighed 12.4 kg and consumed 8 W. In contrast, for AIRS, the figures are 156 kg and 256 W, respectively, while for CrIS they are 81 kg and 91 W. These numbers have major repercussions on costs, not only for the instrument but also for the spacecraft and launch, if flown as an independent mission.

Other features that contribute to a low-cost design are:

- gravity-gradient stabilization is satisfactory for a nadir-viewing instrument;
- there is only one spectral range and one detector type;
- the detectors are uncooled;
- system noise is not an issue for climate averages;
- along-track scanning is not necessary because the field of view is large;
- cross-track scanning does not increase the accuracy of climate averages and
- the FTS contains a minimum number of optical components.

Low cost is essential to maintain climate monitoring when indefinite longevity is required. Some project managers equate low cost to high risk; but the opposite is true if low cost is achieved by means of a disciplined focus on the essential mission objective, which is demonstrable accuracy.

Finally, appropriate technology is not the only requirement for a successful climate monitoring system. The climate science community must invest effort, through committees and workshops, to monitor performance and data continually as a climate mission proceeds, to ensure that reliable information is bequeathed to future generations.

Acknowledgements

We are indebted to Professor Murray Salby for permission to use his Global Cloud Imagery data set, to Dr. Ryoichi Imasu of the center for Climate System research at the university of Tokyo for access to the IMG satellite data, and to Professor John Harries and Dr. Helen Brindley of the Imperial College of Science, Technology and Medicine for permission to refer to unpublished manuscripts. Support has been provided by NASA Grant NAG5-8779, and NOAA Contract 50-SPNA-1-00042.

References

[1] Goody R, Anderson J, Karl T, Miller RB, North G, Simpson J, Stephens G, Washington W. Why monitor the climate? BAMS 2002;83:873–8.
[2] Houghton JT, Meira Filho LG, Callendar BA, Harris N, Kattenberg A, Maskell K, editors. Climate change, 1995: the science of climate change. London: Cambridge University Press, 1995.
[3] Goody R, Haskins R. Calibration of radiances from space. J Climate 1998;11:754–8.
[4] Dykema J, Anderson JG. The experimental foundations for high-accuracy Fourier transform observations for space climate Benchmarks. Appl Opt 2003, to be submitted.
[5] Kirk-Davidoff DB, Goody RM, Anderson JG. Analysis of sampling errors for Climate Monitoring Satellites. J Climate 2003.
[6] Hu H, et al. Temporal and Spatial sampling errors in resolved radiances from space. J Climate 2003, to be submitted.
[7] Hanel RA, Schlachman D, Vanous D. Nimbus 4 Michelson interferometer. Appl Opt 1971;10:1376–82.

[8] Nicholas JV. On the thermodynamic accuracy of the ITS-90: Platinum resistance thermometry below 273 K. Metrologia 1995;32(2):71–7.

[9] Rusby RL, Hudson RP, Durieux M, Schooley JF, Steur PPM, Swenson CA. The thermodynamic basis of ITS-90. Metrologia 1991;28(1):9–18.

[10] Mason IM, Sheather PH, Bowles JA, Davies G. Blackbody calibration sources of high accuracy for a spacebourne infrared instrument: the Along Track Scanning Radiometer. Appl Opt 1996;35:629–39.

[11] Fowler JB. A 3rd-generation water bath based black body source. J Res Nat Inst Stand 1995;100(5):591–9.

[12] Best F, Revercomb H, Laporte D, Knuteson R, Smith W. Accurately calibrated airborne and ground-based Fourier transform spectrometers II: HIS and AERI calibration techniques, traceability and testing. Proceedings of the Council for Optical Radiation Measurements (CORM) 1997 Annual Meeting, National Institute of Standards and Technology (NIST), Gaithersberg, MD, 1997.

[13] Revercomb HE, Best FA, Dedecker RG, Dirkx TP, Herbsleb RA, Knuteson RO, Short JF, Smith WL. Atmospheric Emitted Radiance Interferometer (AERI) for ARM. Proceedings of the Fourth Symposium on Global Change Studies, Anaheim CA, American Meteorological Society, 1993.

[14] Revercomb HE, Best FA, LaPorte D, Knuteson RO, Smith WL, Ciganovich N, Dedecker R, Dirkx T, Garcia R, Herbsleb R, Short J. Accurately calibrated airborne and ground-based Fourier transform spectrometers: HIS and AERI instrument design, performance, and applications for meteorology and climate. Proceedings of the Council for Optical radiation Measurements (CORM) 1997 Annual Meeting, National Institute of Standards and Technology (NIST), Gaithersberg, MD, 1997.

[15] Kannenberg R. IR instrument comparison workshop at the Rosenstiel school of marine and atmospheric science (RSMAS). Earth Observer 1998;10:51–4.

[16] Villemaire A, Chamberland M, Giroux J, Lachance RL, Theriault J-M. Radiometric calibration of FT-IR remote sensing instrument. Proc SPIEE 1997;3082:83–91.

[17] Whatmore RW. Pyroelectric devices and Materials. Rep Prog Phys 1986;49(12):1335–86.

[18] Porter SG. A brief guide to pyroelectric detectors. Ferroelectrics 1981;33:193–206.

[19] Keith DW, Anderson JG. Accurate spectrally resolved infrared radiation observation from space: implications for detection of decade-to-century scale climatic change. J Climate 2001;14:979–90.

[20] Anderson JG, Goody RM, Keith DW. Arrhenius: A small satellite for climate research. Harvard University, Cambridge, MA 1996. Proposal to NASA, obtainable from Professor J. Anderson, Department of Chemistry and Chemical Biology, 12 Oxford Street, Cambridge, MA 02138.

[21] Hu H, Strow LL, Keith DW, Anderson JG. Validation of radiative transfer for atmospheric temperature sensing. Proceedings of the 10th Conference on Atmospheric Radiation, Madison WI, American Meteorological Society, 1999.

[22] Dykema JA, Keith DW, Hu H, Lapson L. An accurate radiometer for climate observation. Proceedings of the 10th Conference on Atmospheric Radiation, Madison WI, American Meteorological Society, 1999.

[23] Hu H, Dykema J, Keith D, Lapson L, Anderson J. Intercomparison of atmospheric radiance measurements by two Fourier transform spectrometers flown on the NASA ER-2. In: Smith WL, Timofeyev YM, editors. IRS 2000: Current Problems in Atmospheric Radiation. Hampton VA: Deepak, 2000. p. 1089–92.

[24] Keith DW, Dykema JA, Hu H, Lapson L, Anderson JG. Airborne interferometer for atmospheric emission and solar absorption. Appl Opt 2001;40:5463–73.

[25] Revercomb HE, Buijs H, Howell HB, Laporte DD, Smith WL, Sromovsky LA. Radiometric calibration of IR Fourier-transform spectrometers: Solution to a problem with the high resolution interferometer sounder. Appl Opt 1988;27(15):3210–8.

[26] Mutlow CT, Závody AM, Barton IJ, Llewellyn-Jones DT. Sea-surface temperature measurements by the along-track scanning radiometer on the ERS1 satellite: Early results. J Geophys Res 1994;99:22575–88.

[27] Dykema JA, Keith DW, Anderson JG. The error budget of a benchmark climate radiometer. Appl Opt 2003, submitted.

[28] Jarecke PJ, Folkman MA, Hedman TR, Frink ME. Clouds and the Earth's radiant energy system (CERES): Long-wave calibration plan and radiometric test model (RTM) calibration results. Metrologia 1993;30:223–30.

[29] Aumann HH, Overoye K. The Atmospheric Infrared Sounder (AIRS) on the Earth Observing System: In-orbit radiometric spectral calibration. Proceedings of the 10th Annual International AeroSense Symposium (SPIE), Orlando FL, 1996.
[30] Stumpf KD, Overbeck JA. CrIS optical system design. Proc SPIE 2002;4486:437–44.
[31] Abrams MC, Toon GC, Shindler RA. Practical example of the correction of Fourier-transform spectra for detector nonlinearity. Appl Opt 1994;33:6307–14.
[32] Guelachvili G. Distortion free interferograms in Fourier-transform spectroscopy with nonlinear detectors. Appl Opt 1986;25:4644–8.
[33] Leroy SS. The effects of orbital precession on remote climate monitoring. J Climate 2001;40:4330–7.
[34] Salby ML, Callaghan P. Sampling error in climate properties derived from satellite measurements: Consequences of under sampled diurnal variability. J Climate 1997;10:18–35.
[35] Kobyashi H, Shimota A, Kondo K, Okumara E, Yoshihiko K, Shimoda H, Ogawa T. Development and evaluation of the interferometric monitor for greenhouse gases: A high-throughput Fourier-transform infrared radiometer for nadir Earth observation. Appl Opt 1999;38:6801–7.
[36] Haskins R, Goody R, Chen L. Radiance covariance and climate models. J Climate 1999;12:1409–22.
[37] Brindley HE, Harries JE. The impact of instrument field of view on measurements of cloudy sky spectral radiances from space: application to IRIS and IMG. JQSRT 2003, submitted.
[38] Harries JE, Brindley HE, Sagoo PJ, Bantges RJ. Increases in greenhouse forcing from the Earth's outgoing longwave spectra in 1970 and 1997. Nature 2001;410:355–7.
[39] Steel WH. Interferometers without Collimation for Fourier Spectroscopy. JOSA 1964;45:151–6.
[40] Brindley HE, Harries JE. Observations of the Infrared outgoing spectrum of Earth from space: The effects of temporal and spatial sampling. J Climate 2003, submitted.
[41] Haskins R, Goody R, Chen L. A statistical method for testing a general circulation model with spectrally-resolved data. J Geophys Res 1997;102:16563–81.
[42] Haskins R, Goody R, Chen L. Radiance Covariance and Climate Models. J Climate 1999;12:1409–22.
[43] Goody R, Anderson J, North G. Testing climate models: An approach. BAMS 1998;79:2541–9.