Evaluation of Long-Term Calibrations of the AVHRR Visible Radiances

WILLIAM B. ROSSOW

NOAA/CREST, City College of the City University of New York, New York, New York

JOSEPH FERRIER

Trinnovim LLC, Arlington, Virginia, and NASA Goddard Institute for Space Studies, New York, New York

(Manuscript received 23 July 2014, in final form 28 January 2015)

ABSTRACT

Two systematic calibrations have been compiled for the visible radiances measured by the series of AVHRR instruments flown on the NOAA operational polar weather satellites: one by the International Satellite Cloud Climatology Project (ISCCP), anchored on NASA ER-2 underflights in the 1980s and early 1990s and covering the period 1981–2009, and one by the PATMOS-x project, anchored on comparisons to the MODIS instruments on the *Aqua* and *Terra* satellites in the 2000s and covering the period 1979–2010 (this result also includes calibration for the near-IR channels). Both methods have had to extend their anchor calibrations over a long series of instruments using different vicarious approaches, so a comparison provides an opportunity to evaluate how well this extension works by cross-checking the results at the anchor points. The basic result of this comparison is that for the ''afternoon'' series of AVHRRs, the calibrations agree to within their mutual uncertainties. However, this retrospective evaluation also shows that the representation of the time variations can be simplified. The ISCCP procedure had much more difficulty extending the calibration to the ''morning'' series of AVHRRs with the calibrations for *NOAA-15* and *NOAA-17* exceeding the estimated uncertainties. Given the general agreement, a new calibration for all AVHRR visible radiances (except TIROS-N, *NOAA-6*, *NOAA-19*, and *MetOp-A*) is proposed that is based on the average of the best linear fits to the two time records. The estimated uncertainty of these calibrations is ±3% absolute (scaled radiance units).

1. Introduction

Weather satellite imaging instruments were first launched in the 1960s and became fully and continuously operational in the early 1970s, first on sun-synchronous polar orbiters, providing global coverage at 12-h intervals (later 6 h with two operational satellites), and then additionally on geostationary satellites, providing regional coverage but at much shorter time intervals ranging from 30 to 180 min (later 5–15 min). These instruments were intended to provide sequential images of cloud patterns indicative of evolving weather systems. Hence, the two wavelengths selected for these images were the ones at which the atmosphere is most transparent and the contrast of cloudy and clear scenes is largest: visible (0.5–0.7 μm) for daytime imagery and the ''window'' infrared (10–12 μm) for all-day imagery. Measurements at these wavelengths were also technologically easier because the radiances coming from Earth are relatively large and the available sensor sensitivity was more than adequate. Since these data were not specifically intended to study clouds or cloud processes, radiance calibration was not a major concern, although early uses were made to study cloud variability (e.g., Zangvil 1975) and even to develop ''climate'' statistics, such as the mean geographic distribution and seasonal variations of cloudiness (e.g., Arking 1964; Reynolds and Vonder Haar 1977). However, in the late 1970s, growing concern about the possibility of climate change induced by human activities drew attention to the need to quantify cloud-climate feedbacks for which global information about cloud physical properties and processes was lacking. As no experimental satellite instruments, designed specifically to obtain cloud information, were planned—indeed, none were flown until this past decade—this cloud information would have to come from the exploitation of the weather satellite measurements, which required radiance calibrations.

Corresponding author address: William B. Rossow, CREST, City College of the City University of New York, T-107, Steinman Hall, 140th Street and Convent Avenue, New York, NY 10031.
E-mail: wbrossow@ccny.cuny.edu

DOI: 10.1175/JTECH-D-14-00134.1

Unlike the infrared channels on the operational imagers (and later infrared temperature–humidity sounders), the solar wavelength channels on operational satellites, even to this day, usually do not have onboard calibration capability, so the earliest calibrations were obtained from so-called vicarious approaches that used the measured reflectivity of ''known'' targets (e.g., Che and Price 1992; Price 1987; Frouin and Gautier 1987; Slater et al. 1987; Staylor 1990; Vermote and Kaufman 1995). For a time in the 1980s and early 1990s, beginning with *NOAA-7* and continuing through *NOAA-11*, NASA and NOAA cooperated to conduct ER-2 aircraft underflights of the polar-orbiting satellites with laboratory-calibrated spectrometers (Hovis et al. 1985). Kriebel (1981) performed a similar calibration for the early Meteosat imagers. When the International Satellite Cloud Climatology Project (ISCCP) began in 1982, the primary goal was to quantify the main radiative properties of clouds and their diurnal, synoptic, seasonal, and interannual variability. To this end, the operational calibration of *NOAA-7*, which appeared to be adequately stable and accurate (compared to what was known at that time about the reflectances of various land surface types; cf. Matthews and Rossow 1987; Rossow and Lacis 1990), was taken as the initial calibration standard (Brest and Rossow 1992) to which all the geostationary radiometers were normalized (Desormeaux et al. 1993). However, after less than two years of data collection, *NOAA-7* was replaced by *NOAA-9*, requiring development of some process to transfer the calibration standard to successive radiometers and to monitor the stability of each radiometer over its lifetime (Brest et al. 1997). Two subsequent cooperative studies led by NASA (Whitlock et al. 1990) and NOAA (Nagaraja Rao and Chen 1995) evaluated a number of vicarious results and combined these with the most comprehensive set of underflights for any radiometer to establish the *NOAA-9* AVHRR as the best calibration standard available. The approach for transferring the calibration from one instrument to the next using time-overlapped observations had to be replaced with a different procedure when *NOAA-11* failed unexpectedly and was not replaced by *NOAA-14* until five months later (subsequently, no overlapping data were provided for *NOAA-14* through *NOAA-16*, *NOAA-16* to *NOAA-18*, and *NOAA-18* and *NOAA-19*). The new procedure was also necessitated by the interruption of the *NOAA-11* time record by the Mount Pinatubo volcano, which changed Earth's target values. The new procedure that has now been applied to the whole time record compares the scene reflectances for each AVHRR from a large number of Earth targets, together comprising complete global coverage, and adjusts these results to match the *NOAA-9* climatology of the same target reflectances.

Originally planned to last five years from July 1983, ISCCP continues today 30 years later and has provided until now the only source of visible calibration information for the whole weather satellite constellation. Some issues concerning the overall accuracy of the ISCCP calibration, even for the series of AVHRRs, were left unresolved: 1) The short record available for the *NOAA-7* AVHRR made determination of any sensor degradation uncertain because of the possible and uncertain effects of anisotropic reflection from the land surface targets as the satellite orbit drifted to overpass times later in the day (larger solar zenith angles). 2) The calibration of earlier morning satellites (*NOAA-8*, *-10*, *-12*, *-15*) was generally more uncertain because their early morning overflight times resulted in more extreme sun geometry that produced relatively low radiance values, even for brighter targets. 3) Because of the orbit drifts, overlapping observations from two afternoon polar orbiters in the series occurred at substantially different times of day (different solar zenith angles) that made the transfer of calibration uncertain. The estimated relative uncertainty of the ISCCP AVHRR calibration record is estimated to be about ±2%–3%; combined with the absolute uncertainty of the ER-2 anchor value, the absolute uncertainty is ±7% for radiances (Brest et al. 1997).

The main question about the ISCCP calibration record now is whether it still maintains its earlier accuracy more than 20 years beyond its anchor point in October 1986. The record is maintained by assuming the reflectance of Earth, as a statistical aggregate of many targets, is constant (except for the El Chichón and Mount Pinatubo periods). Since the ER-2 underflights were discontinued in the early 1990s, there has been no independent verification of the ISCCP calibration for later satellites. The importance of this question is raised by the interpretation of the long-term variations in the ISCCP cloud cover and optical thickness, as the former varies and the latter remains nearly constant except for a small increase starting in the 2000s. When used together to calculate radiative fluxes, this record reproduces a small variation of Earth's albedo that is quantitatively similar to that inferred from the long ERBS record (Zhang et al. 2004).

In the past decade, experimental imaging instruments have been flown with much more attention paid to radiance calibration, including better prelaunch instrument characterization and calibration traceable to international standards, (sometimes multiple) onboard calibration measurements, and in-orbit procedures to monitor calibration stability. For instance, all of these features apply to the calibration of the MODIS instruments, which are estimated to have an absolute

reflectance uncertainty of ±2%, equivalent to an absolute radiance uncertainty of ±5% (Xiong et al. 2007, 2010). Key aspects of the MODIS calibration are that the solar diffusor, used to monitor the calibration of MODIS, is itself monitored by a separate instrument and that the stability of the calibration is also checked by viewing the moon about 10 times per year. The extensive knowledge of the MODIS calibration, especially the high quality for *Aqua* MODIS (Xiong et al. 2010; Wu et al. 2013), provides an opportunity to verify the ISCCP calibration of AVHRR in the more recent epoch.

Heidinger et al. (2010) and Molling et al. (2010) have performed a detailed cross calibration of recent AVHRRs (*NOAA-14* through *NOAA-19* and *MetOp-A*) directly to MODIS on *Terra* and *Aqua* and have extended these results to the earlier AVHRRs (TIROS-N through *NOAA-12*) by vicarious measurements of two standard Earth targets and some coincident measurements between pairs of instruments. This second AVHRR calibration record (ISCCP did not calibrate TIROS-N or *NOAA-6*) is anchored on the current-day MODIS calibrations but extended back in time by assuming the constancy of the two targets. The estimated absolute reflectance uncertainty of this result, which we will call the PATMOS-x product, is about ±2%–3%, which is equivalent to a radiance uncertainty of about ±7%.

Thus, we have two independent absolute calibrations of the AVHRR series, one anchored on the ER-2 underflights of *NOAA-9* in 1986 and one anchored on MODIS in the 2000s. Both sets of calibration coefficients are extended from their anchor points by more than 20 years using Earth targets that are assumed to be constant over time. Experience from these two studies (ISCCP and PATMOS-x) suggests that Earth is at least more constant than the radiometer calibrations (Brest and Rossow 1992; Brest et al. 1997; Heidinger et al. 2010). The purpose of the current study is to compare the ISCCP-ER-2 and PATMOS-x-MODIS calibration records to cross verify them and to tie together the two anchoring absolute calibration references into a single calibration record. In addition, a detailed comparison will evaluate the limitations of these two methods. For ISCCP, only very short records were available for *NOAA-8* and *NOAA-15*. Moreover, the ISCCP approach to normalizing the other morning AVHRRs (*NOAA-10*, *-12*, *-17*) is known to be more uncertain. The robustness of the assumption of constant target reflectances to extend the record needs to be evaluated for both methods: the ISCCP method employs a much larger number and a variety of targets, comprising the whole Earth, whereas the Heidinger approach uses only two small targets, a desert area and an ice sheet.

TABLE 1. Periods of operation of NOAA-series AVHRRs and period used for ISCCP to date.

Satellite	Orbit	Operational period	Imaging data for ISCCP
TIROS-N	p.m.	Oct 1978–Jan 1980	Not used
NOAA-6	a.m.	Jun 1979–Nov 1986	Not used
NOAA-7	p.m.	Aug 1981–Jun 1986	Jul 1983–Jan 1985
NOAA-8	a.m.	May 1983–Oct 1985	Oct 1983–Jun 1984
NOAA-9	p.m.	Feb 1985–May 1994	Feb 1985–Oct 1988
NOAA-10	a.m.	Nov 1986–Sep 1991	Dec 1986–Aug 1991
NOAA-11	p.m.	Nov 1988–Sep 1994	Nov 1988–Aug 1994
NOAA-12	a.m.	May 1991–Dec 1998	Sep 1991–Dec 1998
NOAA-14	p.m.	Dec 1994–May 2007	Feb 1995–Sep 2001
NOAA-15	a.m.	May 1998–	Jan 1999–Jul 2000
NOAA-16	p.m.	Sep 2000–	Oct 2001–Dec 2005
NOAA-17	a.m.	Jun 2002–Apr 2013	Jul 2002–Dec 2009
NOAA-18	p.m.	Aug 2005–	Jan 2006–Dec 2009
MetOp-A	a.m.	Jun 2007–	Not used yet
NOAA-19	p.m.	Jun 2009–	Not used yet

Section 2 provides some basic background information about the NOAA polar-orbiting satellites and their AVHRR instruments, while section 3 provides more details about the two calibration procedures, a description of an additional calibration procedure using optically thick cloud targets and the open ocean to independently evaluate both calibrations, and some information about the MODIS calibration. Section 4 shows the comparison of results, and section 5 summarizes the situation for obtaining the visible channel calibration of the whole AVHRR series and proposes a merged result. Section 6 discusses some key aspects of the interpretation of these results.

2. Some additional background on AVHRR

Most of the background information about the AVHRR series can be found in Molling et al. (2010), including the date range covered by each satellite: Table 1 shows the operational date range and the range covered by the ISCCP data collection. Table 2 compares the spectral responses of the different AVHRRs in the series by giving the wavelength at peak response, the bandwidth (half-power wavelength differences), the average wavelength (no weighting), the average wavelength (weighted by the solar spectrum), and the instrument solar constant (the actual spectral response functions can be obtained at http://isccp.giss.nasa.gov/docs/response.html). The solar spectrum used for Table 2 is from Thuillier et al. (2003); the original values used by ISCCP were based on Neckel and Labs (1984), which differ little from the newer tabulation in the wavelength range of these channels.

Some understanding of the behavior of these radiometers is suggested by noting which ones were built under the same contract, as it was the usual practice to build all of the radiometers in a group much more

TABLE 2. Spectral response functions for AVHRR visible channels in terms of wavelength (λ, μm) at peak response, bandwidth (half-power-wavelength difference, μm), average wavelength (μm), the solar-spectrum-weighted average wavelength (μm), and the instrument solar constant (watts m^{-2} sr^{-1}). The first value for the solar constants are the ones corresponding to the ISCCP calibration discussed here (Neckel and Labs 1984); the values in parentheses are slightly revised values to be used with the revised calibration (Thuillier et al. 2003).

	Peak λ	Bandwidth	Avg λ	Weighted-avg λ	Solar constant
NOAA-7	0.630	0.115	0.633	0.630	56.66 (56.688)
NOAA-8	0.670	0.116	0.634	0.631	56.70 (56.722)
NOAA-9	0.680	0.130	0.638	0.634	60.91 (60.954)
NOAA-10	0.660	0.113	0.631	0.628	56.89 (56.920)
NOAA-11	0.680	0.126	0.639	0.635	58.02 (58.648)
NOAA-12	0.680	0.128	0.643	0.638	63.43 (63.900)
NOAA-14	0.680	0.133	0.642	0.639	65.42 (65.464)
NOAA-15	0.642	0.092	0.634	0.632	44.70 (44.740)
NOAA-16	0.646	0.092	0.634	0.632	42.54 (42.514)
NOAA-17	0.648	0.090	0.636	0.634	43.36 (43.365)
NOAA-18	0.648	0.090	0.637	0.635	42.34 (42.327)

rapidly than actually needed and store them until needed. *NOAA-8* and *-10* (also *NOAA-6*) were four-channel models that did not have the so-called split-window channel at about 12-μm wavelength. Consequently, *NOAA-10* was in storage for a number of years before being launched. The next batch of five-channel instruments (with the addition of the 12-μm channel) included *NOAA-7*, *-9*, *-11*, *-12*, and *-14*. Thus, *NOAA-12* and *-14* were in storage for many years by the time of launch. The final batch of six-channel instruments (a new channel at 1.6 μm was added but data were collected from this channel only during the daylight portion of the orbit and only occasionally) includes *NOAA-15* through *NOAA-19*. This last series also introduced a ''bilinear'' counts-to-radiance relationship for the visible channel.

TABLE 3. Large changes of nominal (original NOAA) calibration or absolute (ISCCP) calibration.

Satellite	Date	Change
NOAA-7		None
NOAA-8		None
NOAA-9		None
NOAA-10	May 1989	Nominal gain and offset changed by NOAA after degassing procedure
NOAA-11	Jan 1989	Absolute gain changed by ISCCP
	Sep 1990	Nominal gain and offset changed by NOAA
	Jan 1994	Absolute gain changed by ISCCP
	Aug 1994	Absolute gain changed by ISCCP
NOAA-12	May 1994	Nominal gain and offset changed by NOAA
	Sep 1994	Absolute gain changed by ISCCP
	Nov 1994	Absolute gain changed by ISCCP
NOAA-14		None
NOAA-15		None
NOAA-16		None
NOAA-17		None
NOAA-18	Jul 2007	Normalized offset changed by ISCCP

Table 3 summarizes some notable changes made in the calibrations reported by NOAA (referred to as nominal or prelaunch calibration) or made by ISCCP.

Some further detail about orbit geometry constraints on solar illumination is summarized in Table 4, which shows the mode values of the distribution of the values ($\geq$0.4) of the cosine of the solar zenith angle μ_0 associated with tropical deep convective clouds at the beginning, in the middle, and at the end of the life of each satellite. When the mode value is given as 0.4, it means that the orbit has drifted close enough to or past the terminator, so that many fewer observations of the tropics are available with adequate solar illumination. For ISCCP the key issue is the range of radiances provided by the targets, which depends on the combination of their reflectances and the solar zenith angles at which they can be viewed. For PATMOS-x the key issue is how often the two targets used, the Libyan desert and Dome C in Antarctica, can be viewed under suitable illumination conditions.

TABLE 4. Mode value of cosine of solar zenith angles $\geq$ 0.4 for all DCC, defined by infrared brightness temperatures $<$ 210 K, at the beginning, middle, and end of record for each AVHRR (see Table 1 for data record period for each satellite). If the mode value = 0.4, the satellite has drifted too close to the terminator for reliable results.

	Beginning	Middle	End
NOAA-7	0.69	0.58	0.51
NOAA-8	0.44	0.41	0.47
NOAA-9	0.84	0.70	0.41
NOAA-10	0.44	0.40	0.41
NOAA-11	0.80	0.70	0.40
NOAA-12	0.43	0.40	—
NOAA-14	0.92	0.73	0.40
NOAA-15	0.40	0.40	0.40
NOAA-16	0.80	0.72	0.66
NOAA-17	0.88	0.91	0.69
NOAA-18	0.90	0.92	0.84

3. Calibration methodologies

a. ISCCP methodology

The ISCCP procedure determines a gain and offset value, so that the scaled radiance L^* is given by

$$L^* = \mathrm{GI} \times \mathrm{CT}_8 + \mathrm{OI}, \quad (1)$$

where GI is the ISCCP gain, CT_8 is the 8-bit radiance count value, and OI is the ISCCP offset. Both GI and OI are assumed to vary with time; over the lifetime of each satellite, this variation is usually linear unless monitoring evidence suggests otherwise (as it does in some cases). Scaled radiance is the radiance in physical units, such as watt per steradian per square meter, divided by the instrument ''solar constant,'' which is the wavelength integral of the product of the solar spectrum (in the same units) and the normalized (peak response is unity) instrument response function. Thus, scaled radiance is dimensionless. Reflectance is equal to the scaled radiance divided by the cosine of the solar zenith angle and corrected for the deviation of the sun–earth distance from its annual mean.

During the first phase of ISCCP from 1983 through 1990, the nominal calibration of the *NOAA-7* AVHRR was taken as the reference standard and this calibration was initially assumed to be stable. When it was replaced by *NOAA-9* in 1985, evaluation of initial data from *NOAA-9* suggested that the performance of *NOAA-7* had in fact degraded, so a procedure was developed both to monitor the calibration of each radiometer over its lifetime and to transfer the standard to successive radiometers (Brest and Rossow 1992). The monitoring procedure relied on same-month-in-a-year-to-year regressions (e.g., January–January, February–February, etc.) of collocated (all pixels mapped to a 25-km equal-area grid) ''clear sky'' reflectances covering the whole globe with a crude statistical cloud-clearing method that focused on the mode values of the reflectance distributions. A similar approach was used to transfer the afternoon AVHRR standard to the concurrent morning instruments and to transfer the standard from one afternoon satellite to the next in sequence. The sample size for all of these regressions was very large because this procedure used a whole month of data (or a few weeks for the overlapping data) covering the whole sunlit globe at 25-km intervals sorted into 28 geographic targets separated by surface type or vegetation cover. These procedures were applied to the AVHRRs on *NOAA-7* through *NOAA-10* and the first 2.5 years of *NOAA-11* (Brest and Rossow 1992).

Subsequent cooperative studies led by NASA (Whitlock et al. 1990) and NOAA (Nagaraja Rao et al. 1993) evaluated a number of vicarious results, including those from ISCCP, together with the most comprehensive results of ER-2 underflights. These results led to the conclusion that the *NOAA-9* AVHRR was the best calibration standard available, which was adopted by ISCCP (Brest et al. 1997). The six NASA ER-2 flights under *NOAA-9* spanned nearly its complete lifetime and included one set of three measurements taken in just a few days in October 1986 (the only such set of results ever obtained). The absolute uncertainty of the calibration from an individual flight was estimated to be about ±9% (B. Gunther 1988, private communication), so the aggregate absolute uncertainty for the *NOAA-9* calibration in October 1986 is estimated to be slightly more than ±5%. The various estimates of the degradation of this radiometer over its lifetime, together with the ER-2 measurements on four dates, were in excellent quantitative agreement (Nagaraja Rao et al. 1993), so the overall absolute calibration uncertainty for *NOAA-9* visible radiances was taken to be 10% (±5%).

The procedures for monitoring and transferring of calibration had to be replaced by a new approach because of two events. The eruption of Mount Pinatubo in June 1991 during the operations of *NOAA-11* put enough aerosol into the stratosphere to significantly change the earth as a standard, stable target; thus, the monitoring of the AVHRR calibration had to be suspended until the aerosol cleared and enough statistics on the other side of the time gap (from fall 1991 through 1992) could be collected. But then *NOAA-11* failed unexpectedly in August 1994 and *NOAA-13* failed after launch, so there was no afternoon AVHRR available until February 1995, when *NOAA-14* began operations. The new and current procedure now compares the scene reflectances for a large number of Earth targets, together comprising complete global coverage, obtained each month over the lifetime of each AVHRR against the time record from the *NOAA-9* climatology and adjusts the calibration to produce statistical agreement: although ocean and cloud reflectances are also examined, the adjustment is based on the aggregate statistics of the land surface. In other words, the *NOAA-9* record provides a climatology of global land target reflectances and the measurements of previous and subsequent AVHRRs are adjusted to match, equivalent to assuming that the statistical aggregate of target reflectances over Earth is constant in time. The period affected by the Mount Pinatubo aerosols is not included in the analysis of *NOAA-11*. Although monthly statistics are obtained for each satellite, the record of aggregate land surface reflectances over the whole lifetime of each satellite is fit with a straight line to remove sensor trends and then the whole record is adjusted to the *NOAA-9* climatology.

Hence, we interpret the very small month-to-month variations as statistical variations and/or variations of some of the targets, not calibration variations.

Using the global distribution of reflectances, including some ocean values (reduced population to equalize statistical weight) and the permanent ice sheets, provides an approximately uniform and very robust sampling of the whole visible reflectance range. The final values of GI and OI for a particular AVHRR are obtained in two steps: normalization of the nominal (prelaunch) calibration to the *NOAA-9* target reflectance climatology and then the correction for slow time variations (usually a linear trend) over the lifetime of the satellite based on monthly comparisons to the *NOAA-9* target climatology. Note that since the nominal calibration is used to produce the initial scaled radiances, the bilinear calibrations are properly used; the gain and offset adjustments that are determined from the normalization and detrending steps are then applied equally to both ranges of the bilinear gain. This is equivalent to assuming that the shape of the bilinear response is preserved. The relative uncertainty of normalizing all other radiometers to *NOAA-9* by this procedure is estimated to be about 3%–5% (Brest et al. 1997). The new procedure was retroactively applied to all of the older data to produce a consistent calibration over the whole AVHRR record (see summary figures in section 5). The ISCCP calibration is provided for every month and has an estimated total absolute radiance uncertainty of ±7%.

b. *PATMOS-x methodology*

The PATMOS-x procedure determines a gain with the space count values obtained directly from the global area coverage (GAC)-format data (the space view measurement is obtained on every scan but only the average is reported for every 50 scan lines), so that the scaled radiance L^* is given by

$$L^* = \mathrm{GP}(\mathrm{CT}_{10} - \mathrm{CT}_{\mathrm{sp}}), \tag{2}$$

where GP is the PATMOS-x gain, CT_{10} is the 10-bit radiance count value, and CTsp is the 10-bit space count. For the bilinear AVHRRs, a single set of CT_{10} values is produced by scaling the count interval by 0.5 below the ''break'' count and by 1.5 above this (Heidinger et al. 2010). The values of GP are assumed to vary quadratically with time, and CTsp is observed to be nearly constant over time for each satellite. The values used for CTsp in the PATMOS-x analysis are not documented in Heidinger et al. (2010) or contained in the available database, but average values can be obtained from Ignatov et al. (2005).

Heidinger et al. (2010) carry out a thorough calibration of both solar channels on the AVHRR series (ISCCP only calibrates the visible channel) by comparison to MODIS measurements (the visible band is number 1 with a central wavelength of 0.645 μm) from both *Terra* and *Aqua*, which are assumed to be the calibration standard. The whole AVHRR series is then calibrated in three parts. In the first part, AVHRR and MODIS measurements, both averaged in a 0.5° equal-angle mapping, are directly matched in a simultaneous nadir overpass (SNO) method to determine the calibration of the AVHRRs on *NOAA-14* through *NOAA-19* and *MetOp-A*. Small differences in spectral responses are accounted for using an empirical fit to radiative transfer model simulations of scene reflectances for a variety of atmospheric water vapor amounts, specified aerosols, ocean surfaces with and without sea ice, and clouds. The linear regression of the MODIS scene reflectances to AVHRR count values for each time is forced through zero. These results are obtained only in the Northern Hemisphere for July in each year (there are only two determinations for *NOAA-14*). In the second part, the MODIS measurements are used to characterize the cloud-free 0.5° scene reflectances for two targets, Dome C in Antarctica (during December–February) and the Libyan desert (only used for afternoon polar orbiters). Again, radiative model simulations of scene radiances are used to develop an empirical fit that accounts for spectral differences between MODIS and AVHRR. These results are used to evaluate the SNO results for AVHRR compared to MODIS but in the final analysis, both the SNO and the more numerous target results are mixed together with weighting reflecting the estimated uncertainties. The target reflectances are assumed to be constant over the whole AVHRR epoch (1978–2009) and then used, in the third part, to calibrate the remaining afternoon AVHRRs. Because of the more extreme solar illumination of the morning polar orbiters, the calibration is transferred to them using a combination of the Dome C comparisons and SNO comparisons between concurrent morning and afternoon orbiters (only for July after 1986).

The combination of methods applied—SNO once per year, Dome C seasonal, Libyan desert every month—provides nearly monthly time sampling of calibrations for all of the AVHRRs flown to date. The PATMOS-x approach is equivalent to assuming that the AVHRRs do not undergo any transient changes in calibration on time scales less than one year or any changes in their (quadratic) degradation rate over the lifetime of each instrument. Heidinger et al. (2010) adopt the claimed uncertainty for the MODIS calibration of ±2% for reflectances and assume that this holds for the direct

MODIS-to-AVHRR SNO results. The remaining results depending on the Dome C and Libyan desert target methods are estimated to have uncertainties of ±3% and ±4% for reflectances, respectively. The final results for GP come from the fit of a quadratic-in-time function to the whole record of calibration results for each AVHRR; the scatter of results about these fits is about ±2%, but the overall estimated uncertainty of reflectances is estimated to be about ±4%–5% (about ±7% for radiances). However, it should be noted that the mixture of MODIS SNO and target-based results for the final calibration of the AVHRRs overlapping MODIS in time produces a slightly high bias with respect to the MODIS SNO results of 0.5%–1.5% (see sections 4b and 4c).

c. *"Deep convective cloud" methodology*

As an independent evaluation of the ISCCP and PATMOS-x calibrations, we examined the global distributions of reflectances collected on different time scales from daily to monthly and investigated the time variation records of several percentile reflectances from these distributions. We found that stable statistics were obtained for weekly and longer accumulations over land and ocean areas, separately, with the cosine of the solar zenith angle μ_0 constrained to be ≥ 0.4. The latter constraint also requires monitoring the total number of pixels included in the statistics, which can become too small when the orbit drifts toward the terminator. Seasonal changes in illumination of the ice sheets introduce stronger time variations for the land areas. So, although we routinely examine the reflectance values at several percentiles (10%, 25%, 50%, 90%, 99%) for both land and ocean, we show here only the two most stable reflectance statistics. The low end of the reflectance range is represented by the 10th percentile reflectance values over oceans (called "10th") and the high end of the range is represented by the mode of the distribution of deep convective cloud reflectances. We found that the reflectances even at the 99th percentile were too variable, so we obtained a specific subset of cloudy reflectances, identified by infrared brightness temperatures < 210 K that represent the nearly saturated reflectances of very optically thick clouds. Note that these clouds have similar visible reflectivities to the ice sheets but, being located in the tropics with larger solar insolation, exhibit larger radiances. This so-called deep convective calibration (DCC) method [Doelling et al. (2004) and Hu et al. (2004) for broadband albedos; Sohn et al. (2009) and Ham and Sohn (2010) for visible reflectances] uses the brightness temperature threshold to isolate the optically thickest clouds (cf. Fu et al. 1990). Such clouds exhibit reflectances that are asymptotically approaching the theoretical limit for a conservative scattering medium (there is very, very weak absorption by clouds at visible wavelengths). We use monthly histograms of these reflectances over all locations that exhibit a mode value very close to the maximum value observed; the use of the mode value is much less sensitive to variations of the clouds.

Based on our radiative transfer calculations of the reflectances for "infinitely thick" cloud layers (in this case, optical thickness = 300), we examined the angle dependence of these reflectance values (cf. Sohn et al. 2009). For isolated cloud layers, these values are relatively invariant with viewing zenith, solar zenith, and relative azimuth angle for forward scattering geometry varying by < 0.02 for $0.4 \leq \mu_0 \leq 1.0$: the values actually decrease slightly for μ_0 values above and below about 0.6–0.8 depending on relative azimuth angle. The dependence on μ_0 is stronger for backscattering geometry, decreasing monotonically from about 1.02 at $\mu_0 = 1.0$ to about 0.90 at $\mu_0 = 0.4$. Given the "morning" and "afternoon" orbit geometry, the observations are dominated by forward scattering, but selecting the mode value in the histograms emphasizes the larger reflectances. Taking Rayleigh scattering into account, which is minimal for such high cloud tops, and ozone absorption, the reflectances of such clouds should be in the range 0.8–0.9.

The choice of open ocean (10th percentile) and DCC for evaluating these calibrations is motivated by the fact that the reflectivities of these two objects are very well understood theoretically and cover almost the whole range of measured radiances.

d. *MODIS calibration*

The calibration of the solar wavelength channels of the MODIS instruments on the *Terra* and *Aqua* satellites is performed by comparison to an onboard solar diffuser (SD) that provides a reflectivity standard (Xiong and Barnes 2006; Xiong et al. 2007, 2010), not an absolute radiance standard. The frequency of these calibration measurements has varied from about weekly early in the mission to only about monthly in the extended mission. Note that this standard provides an absolute radiance standard only if values for the spectral solar constant of the channels are determined (from Thuillier et al. 2003 for band 1). The solar diffuser is, in turn, monitored by another onboard instrument, the solar diffuser stability monitor (SDSM), that is a ratioing radiometer consisting of an integration sphere and nine spectral detectors. The SDSM compares alternate measurements of direct sunlight (through a screen to match responses for some channels) and the radiance from the SD, as well as making measurements of space. Finally, the calibration stability is checked by nearly monthly observations of the moon. The scaled radiance is given by

$$L^* = \mathrm{GM} \times \mathrm{CT}^*_{16} , \quad (3)$$

where CT^*_{16} is the 16-bit MODIS count value that has been corrected to remove the effects of instrument background (using space views), instrument temperature, and viewing angle changes. GM is the MODIS gain given by

$$\mathrm{GM} = (\rho \cos\theta \times \Delta)/d^2 \mathrm{CT}^*_{16} , \quad (4)$$

where ρ is the bidirectional reflectance of the SD, determined prelaunch; θ is the solar zenith angle at the SD at the time of the calibration event; Δ is a correction for SD degradation determined from the SDSM and lunar observations; and d^2 is the ratio of the sun–earth distance at the time of observation to the annual mean distance. The values of CT^*_{16} in (4) are obtained when the instrument views the SD. A correction for a diffuser screen is not shown in (4) because the screen is not used for the calibration of band 1.

The reported results indicate that the MODIS (Collection 5) calibration for band 1 meets the requirement of ±2% for reflectances (±5% for radiances), that the degradation of the band 1 channel has been <0.5% decade^{-1} over the lifetime of the *Aqua* MODIS, and that the *Terra* MODIS degradation rate has been ~2.5% decade^{-1} (Wu et al. 2013). The degradations have been corrected in Collection 6.

4. Comparison of results

In what follows we compare gain values and their evolution over time from the two analyses, ISCCP and PATMOS-x (GI and GP, respectively). The offset values (OI and OP, respectively) were found to be nearly identical. Based on these results, we develop a revision of the ISCCP calibration that is applied to the original ISCCP calibration in the form of a multiplicative factor that varies with time t in months:

$$\delta(t) = \delta_0(1 + t \times \partial G/\partial t), \quad (5)$$

where $\partial G/\partial t$ is the rate of change of the gain per month. This approach is appropriate because, as we will show, the offset values available from both sources are essentially the same and making a correction solely with a multiplicative factor preserves these values. This approach also makes it easier to propagate any changes in AVHRR calibration to the geostationary satellite radiometers that have been normalized to the afternoon AVHRRs. This approach is also simpler when dealing with the bilinear count-to-radiance relationships of the latest AVHRRs.

We illustrate the calibration comparisons between the ISCCP and PATMOS-x with three kinds of figures. The ISCCP calibration in these figures is based on the coefficients used to generate the calibration tables and not the tables themselves (there is only one case, *NOAA-15*, where these are not the same because of an error in implementation). The first type of figure compares the gain values (GI, converted to 10-bit units, and GP, both multiplied by 100) and their time variations over the lifetime of each AVHRR. Note the exaggeration of the vertical scale in these figures (as compared with the scale in the second type of figure described next). The PATMOS-x data points from the Libyan and Dome C targets and from AVHRR-SNO and MODIS-SNO (when available) are also shown with the PATMOS-x quadratic fit (blue solid curves) and a best linear fit (red solid lines). The ISCCP results are shown in their actual form (black solid lines) and by a best linear fit to the actual values (black dashed lines); these are the same except where noted. For the purposes of this display and comparison, the ISCCP calibrations for the bilinear AVHRRs are treated in the same way as in Heidinger et al. (2010), which presents a single gain value that is the weighted average of the gains in the two count ranges. In two cases, we show an alternative bilinear fit to the PATMOS-x points (red dashed lines). In these figures increasing gain values over time implies degrading the sensor response over time.

The second type of figure compares the time records of the scaled radiances obtained by applying the ISCCP and the PATMOS-x calibrations (rescaled to 8-bit units) to three constant 8-bit count values (64, 128, 254) representing the whole range of values (since count = 128 is near the bottom of the upper gain range for the bilinear AVHRRs, we also show scaled radiance values for count = 120, which is near the top of the lower gain range for these satellites). Note that the resulting radiance variations over time indicate the changing sensitivity of the sensor, as the target count values in these plots are constant in time. Thus, for example, if the target radiance decreases because the sensor sensitivity degrades, the gain has to be increased over time to correct for this change. In the first type of plot, this is shown directly by increasing gain values over time and in the second type of plot by increasing scaled radiances over time. Since Heidinger et al. (2010) did not report intercept values, because they assumed that the measured space counts contained in the GAC-format data would be available and nearly constant, we construct the scaled radiance for the PATMOS-x calibrations using the preflight measured space count values reported by Ignatov et al. (2005) in his Table 1 for each satellite converted to 8-bit values [Li et al. (2014) have subsequently reported

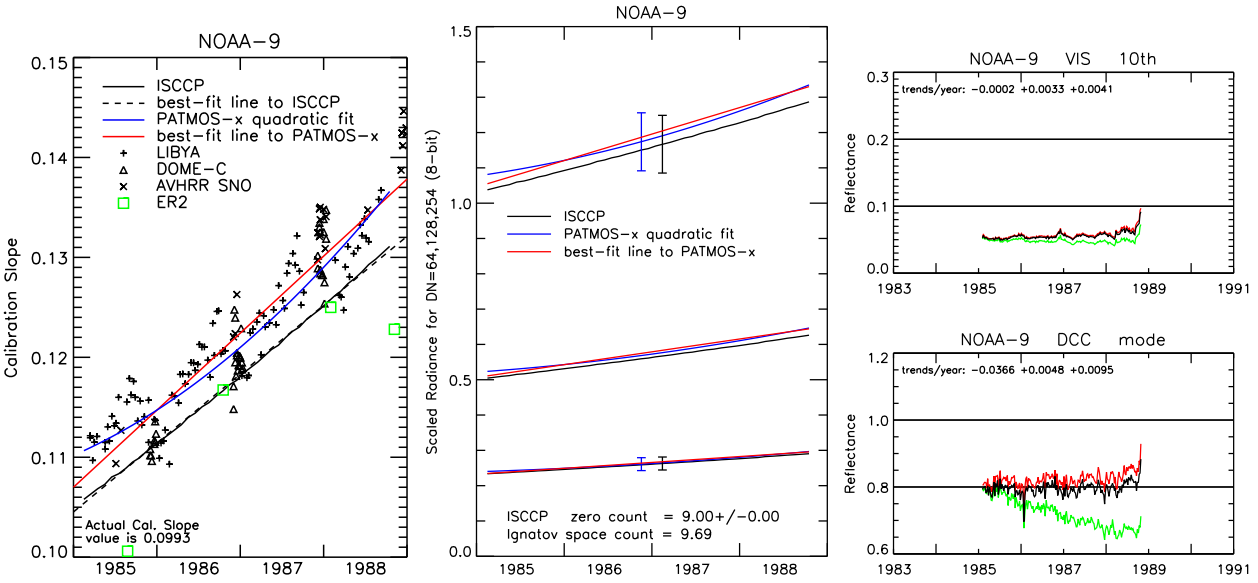


FIG. 1. (a) Comparison of the time variations of gain values (multiplied by 100) to be applied to 10-bit radiance count values for the AVHRR on *NOAA-9*. The individual points from Heidinger et al. (2010) are values obtained by observing the Libyan desert ("+"), Dome C in Antarctica ("Δ"), and SNO observations between *NOAA-9* and the preceding (*NOAA-7*) and the following (*NOAA-11*) afternoon polar orbiters. The green square symbols show the calibration obtained from several ER-2 underflights of this AVHRR. The first point is actually slightly off scale, so its value is shown explicitly; the second point is actually the average of results from three flights and is used as the anchor for the average gain used by ISCCP, where the other points help verify the time variation of the gain. The blue solid curve is the PATMOS-x quadratic fit to these points; the red solid line is the best linear fit to the same points. The black solid line is the actual ISCCP gain values, and the black dashed line is the best linear fit to the ISCCP gain values. (b) Comparison of the scaled radiances obtained by applying the time-varying PATMOS-x calibration gain values [blue solid curve in (a)], and the gains from the best linear fit to the PATMOS-x calibration points [red solid line in (a)] using the indicated space count value from Ignatov et al. (2005), and the actual time-varying ISCCP calibration gain values [black line in (a)] using the indicated space count value to three constant 8-bit count values (64, 128, 254). The error bars indicate the published uncertainty estimates for both calibrations. (c) Comparison of (top) the time variations of 10th percentile visible reflectance values from observations collected over the global oceans for each week with μ_0 constrained to be ≥0.4, and (bottom) the time variations of the mode of the distribution of visible reflectances DCC identified by infrared brightness temperatures <210 K. The green lines show the results for the original calibration for *NOAA-9*, the red lines show the results with the best linear fit to the PATMOS-x calibration points in (a), and the black lines show the results with the ISCCP calibration. The linear reflectance trend per year for the three calibrations (original, ISCCP, PATMOS-x) is indicated.

the values used for PATMOS-x]. These values do not vary much (usually <2 counts rms) over the lifetime of the satellites (Ignatov et al. 2005). The average space count equivalent to the ISCCP offset (OI) and its rms variation over time are also given. In these figures we show the actual ISCCP calibration time record (black solid lines), the PATMOS-x quadratic fit (blue solid curves), and a calibration based on the best linear fit to the PATMOS-x data points (red solid lines) from the first type of figure. We do not necessarily propose to change the ISCCP results to this adjusted version as discussed for each case, but we show this to illustrate the differences.

In the third type of figure, we evaluate the different calibrations in the context of the instrument-lifetime records and whole time series records of reflectance statistics with two statistics. Representing the low end of the radiance range, we show the 10th percentile of the global weekly distribution of reflectances over the whole ocean (called 10th). Representing the high end of the radiance range, we show the mode of the global weekly DCC reflectance distribution over all locations (called DCC; the results over land are the same to better than 1% level). The original calibration is shown by green solid lines, the ISCCP calibration by black solid lines, and the best linear fit to the PATMOS-x results by the red solid lines. These figures show whether the various calibration results have eliminated any significant linear trends over the lifetime of each satellite and over the whole record of afternoon and morning satellites.

a. NOAA-9

NOAA-9 is the anchor for the whole ISCCP calibration record, as it was the only AVHRR instrument that was well and independently calibrated postlaunch by six ER-2 underflights, which are shown as green squares in Fig. 1a (see caption for details). Figure 1a compares the ISCCP and PATMOS-x results for *NOAA-9* by showing the time evolution of the gains, GI (solid black line is the

TABLE 5. Standard deviations (%) between the PATMOS-x set of gain values obtained from various targets and/or satellite-to-satellite comparisons and different least squares fits to these points with all points weighted equally. The values in parentheses for *NOAA-14* show the effects of adopting the weighting of points used by PATMOS-x. The alternate fits shown for *NOAA-10* and *NOAA-14* are bilinear fits as shown in the figures. The ISCCP value is for a linear fit to the ISCCP results but with the mean gain value adjusted to the same gain value as the time average of the PATMOS-x points. The ISCCP values in parentheses for *NOAA-10* through *NOAA-12* show the standard deviations of the actual ISCCP gain values rather than the linear fit.

Satellite	PATMOS-x fit	Best linear fit	Alternate fit	ISCCP
NOAA-7	2.13	2.12	—	3.29
NOAA-8	1.82	2.13	—	2.65
NOAA-9	2.35	2.46	—	2.55
NOAA-10	2.14	3.01	2.13	3.09 (2.69)
NOAA-11	2.29	2.33		2.33 (3.09)
NOAA-12	2.62	2.62	—	2.98 (4.93)
NOAA-14	2.28 (2.28)	2.78 (2.80)	2.21 (2.20)	2.89
NOAA-15	2.75	2.73	—	4.10
NOAA-16	2.24	2.23	—	2.40
NOAA-17	2.51	2.53	—	8.38
NOAA-18	2.38	2.35	—	2.90

actual values and the dashed black line is the best linear fit) and GP (blue curve). That the slope of the best linear fit to the PATMOS-x data points (red solid line) is very similar to that of the ISCCP line indicates very good agreement on the degradation rate, $\partial G/\partial t$, of this AVHRR (Table 5). The small (2%–5%) offset in gain values translates into a difference of scaled radiances at midrange of about 2%–3% (Fig. 1b, where OI is equivalent to an 8-bit space count of 9.00 compared with the average value of 9.69 from Ignatov et al. 2005), well within the absolute uncertainty of these two results. Table 5 compares the original ISCCP gain time evolution to a value determined by a linear fit to the PATMOS-x data points. The SNO results in late 1988 appear to be outliers that may have been produced by the SNO between *NOAA-9* and *NOAA-11*: the former satellite's orbit had drifted near the terminator, so that most daytime observations were at extreme sun angles (mode value of $\mu_0 = 0.41$; Table 4) and the latter satellite appeared to have a rapid decrease of sensitivity shortly after operations began (see section 4c; cf. Brest et al. 1997). These points affect the PATMOS-x quadratic fit but only change the slope of the best linear fit slightly, improving the agreement with the ISCCP results. Figure 1b shows that the best linear fit to the PATMOS-x points actually agrees about as well as the PATMOS-x quadratic fit; Table 5 reinforces this conclusion, showing the scatter of the points is essentially the same for both the linear and quadratic fits. Moreover, previous studies comparing many different results for *NOAA-9* with the ER-2 results did not indicate any such nonlinear behavior (Whitlock et al. 1990) like that shown by the PATMOS-x fit. Nevertheless, all of these detailed differences are smaller than the uncertainties, as illustrated by the error bars in Fig. 1b. Figure 1c shows the 10th and DCC reflectance variations over time: the original calibration shows a trend in DCC of $-3.66\%\ \mathrm{yr}^{-1}$ and almost no trend at 10th, the ISCCP calibration slightly overcorrects to a trend of about $+0.33\%\ \mathrm{yr}^{-1}$ for 10th and $+0.48\%\ \mathrm{yr}^{-1}$ for DCC, and the best linear fit to the PATMOS-x points (excluding the last points) overcorrects to a slightly larger trend of $+0.41\%\ \mathrm{yr}^{-1}$ at 10th and $+0.95\%\ \mathrm{yr}^{-1}$ in DCC. That the ISCCP and PATMOS-x results agree to within their estimated uncertainties provides confirmation of the PATMOS-x extrapolation of the MODIS-based calibration back in time.

b. NOAA-18

NOAA-18 has been directly calibrated against MODIS *Aqua*, and together with *NOAA-16* it forms the anchor for the PATMOS-x calibration record. Figure 2a compares the time evolution of GI (black solid line is actual, black dashed line is best linear fit) and GP (blue curve) for *NOAA-18*. The ISCCP time evolution exhibits a larger linear degradation rate than the best linear fit to the PATMOS-x results (red solid line), which is as good a fit to the PATMOS-x points as the quadratic fit (Table 5). Also note that the MODIS points in Fig. 2a generally lie below the best linear fit and the PATMOS-x quadratic fit: a linear fit to only the MODIS points lowers the gain by about 1%. Despite the difference in gain change over time, the time variation of the OI values actually gives scaled radiances that agree quite well with the PATMOS-x values, except for the largest values (Fig. 2b). Note that the ISCCP offset actually shifts the best agreement for smaller scaled radiances to the beginning of the record (Fig. 2b), even though the gain values differ most then (Fig. 2a): the initial average value of OI was equivalent to a space count of 8.22, less than two counts different from the value reported by

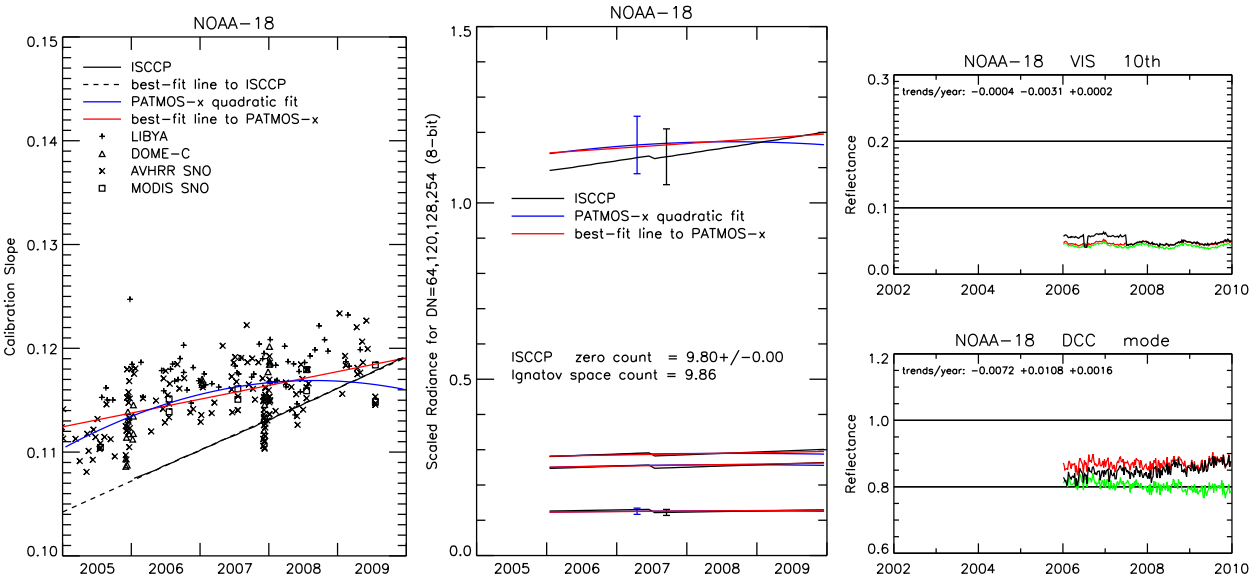


FIG. 2. (a) As in Fig. 1a, but for *NOAA-18*. The AVHRR SNO results are from *NOAA-16* and *NOAA-17*. The "□" symbols show the results from matched MODIS SNO observations. (b) As in Fig. 1b, but for *NOAA-18*. (c) As in Fig. 1c, but for *NOAA-18*.

Ignatov et al. (2005) of 9.86. The scaled radiance differences, even at the end of the record, are <3% at midrange (Fig. 2b), well within the estimated uncertainties. The kink in the ISCCP calibration record (Fig. 2b) that appears in mid-2007 is related to an inadvertent change made to the normalized calibration of this AVHRR (Table 3), where an OI of 0.001 was mistakenly entered as 0.01. The average value of space count after the change is 9.80, which compares better with the average value of 9.86 from Ignatov et al. (2005); this offset error has been corrected in the new calibration (see section 5). Figure 2c confirms that the ISCCP trend correction is slightly too large: the original calibration exhibits a DCC trend (degradation) of $-0.72\%\ \mathrm{yr}^{-1}$, which the ISCCP calibration overcorrects to a positive trend of $+1.08\%\ \mathrm{yr}^{-1}$, whereas the linear fit to PATMOS-x produces a trend of $+0.16\%\ \mathrm{yr}^{-1}$. Although the two calibrations agree well for smaller-scaled radiance values, there is still a difference at the larger values, which is confirmed by the DCC values. Nevertheless, the ISCCP and PATMOS-x results agree to within their uncertainties (error bars in Fig. 2b), which provides confirmation of the ISCCP extrapolation forward in time.

c. *Afternoon satellites (*NOAA-7, NOAA-11, NOAA-14, *and* NOAA-16*)*

NOAA-7 was the first AVHRR examined for degradation by ISCCP; as a consequence, the uncertainties were larger and the corrections more cautious. In particular, the fact that the *NOAA-7* orbit drifted to later in the day over its lifetime (Table 4) meant that the land surface was viewed at systematically larger solar zenith angles and, given that land surface reflectivities are anisotropic (e.g., Verstraete et al. 1990), this implied that the land surfaces should become more reflective, which confounded any degradation estimate. Figure 3a indeed shows that the ISCCP estimate of the degradation rate (actual is black solid line, best linear fit is black dashed line) is smaller than that implied by a linear fit to the PATMOS-x measurements (red solid line; the PATMOS-x quadratic fit, shown by the blue solid curve, is very nearly linear). A hint of this undercorrection can also be seen in Fig. 3c, where the calibration statistics indicate lower values near the end of the *NOAA-7* epoch than at the beginning of the *NOAA-9* epoch. Nevertheless, Fig. 3b shows that the underestimate of scaled radiances at the end of 1984 is only about 3%–4% at midrange, just within the stated relative uncertainty [OI is equivalent to an 8-bit space count of 7.79 as compared to 9.39 from Ignatov et al. (2005)]. The linear and quadratic fits to the PATMOS-x points yield essentially equivalent results with a scatter of about 2%. Figure 3c shows the effect of changing the degradation rate to that given by the best fit to the PATMOS-x data: although the ISCCP calibration made a very tiny improvement at 10th, it hardly reduced the degradation rate in DCC from -0.386% to $-0.363\%\ \mathrm{yr}^{-1}$, whereas the PATMOS-x linear fit reduces the degradation rate to $-0.148\%\ \mathrm{yr}^{-1}$.

NOAA-11 was very difficult to calibrate because of a suspected rapid change just after launch, which may explain the first set of SNO results in Fig. 4a that appear to be outliers, and the occurrence of the Mount Pinatubo volcano 20 months later, so the 70-month record had to

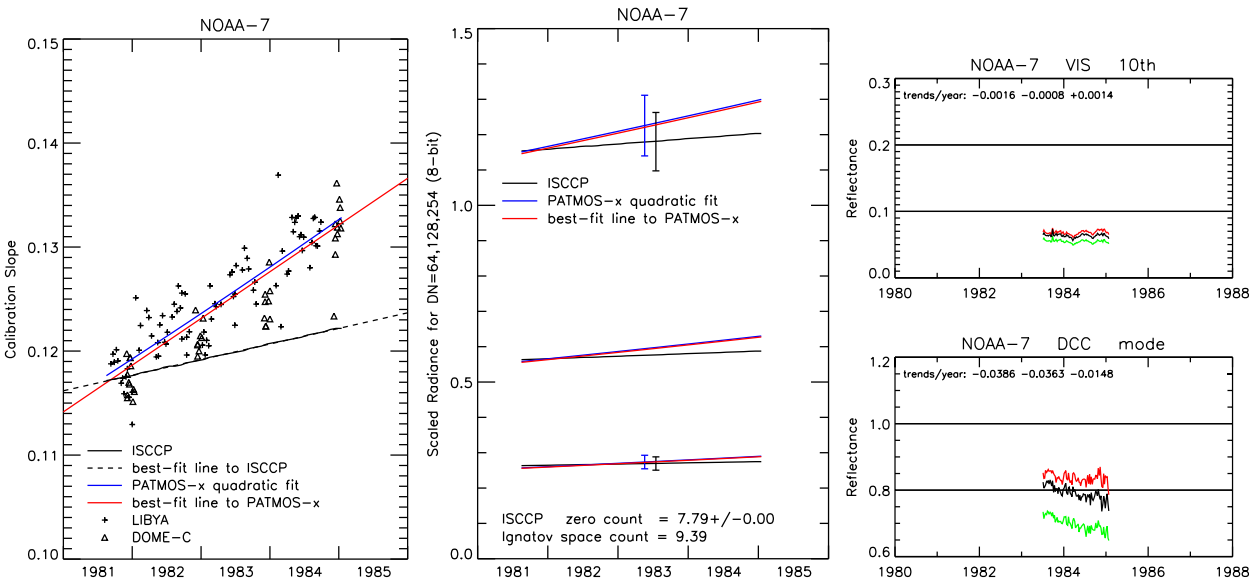


FIG. 3. (a) As in Fig. 1a, but for *NOAA-7*. There are no AVHRR SNO results. (b) As in Fig. 1b, but for *NOAA-7*. (c) As in Fig. 1c, but for *NOAA-7*.

be evaluated as two separate parts of 20 and 32 months with an 18-month gap. Figure 4a compares the PATMOS-x points, its quadratic fit (blue solid curve), a best linear fit to the PATMOS-x points (red solid line), a linear fit to the ISCCP GI time record (dashed black line), and the actual time evolution of GI (black solid line). The actual evolution of GI (and OI) was constructed from a linear fit to monitoring data before Mount Pinatubo (eruption occurred in June 1991) and a linear fit after Mount Pinatubo (1993 onward) extrapolated backward. In addition there appeared to be a more rapid evolution near the end of life that, together with the initial rapid evolution, was accounted for by changes in GI (black solid line in Fig. 4a) and OI (not shown). The sudden change in GI in late 1990 was also necessitated by a change of the nominal calibration by NOAA on 27 September 1990 (Table 3). Nevertheless, Fig. 4a indicates that the linear fit to the ISCCP results captured the same general degradation rate, overall, as PATMOS-x (removing the first SNO results slightly

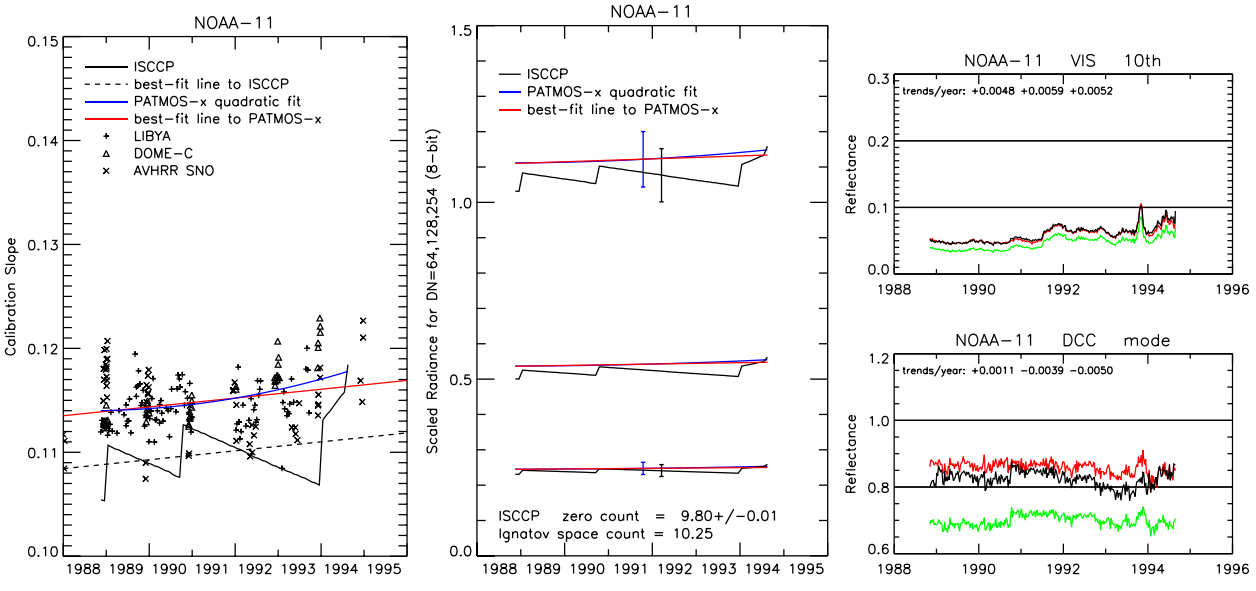


FIG. 4. (a) As in Fig. 1a, but for *NOAA-11*. The AVHRR SNO results are from *NOAA-9* (preceeding) and *NOAA-14* (following). (b) As in Fig. 1b, but for *NOAA-11*. (c) As in Fig. 1c, but for *NOAA-11*.

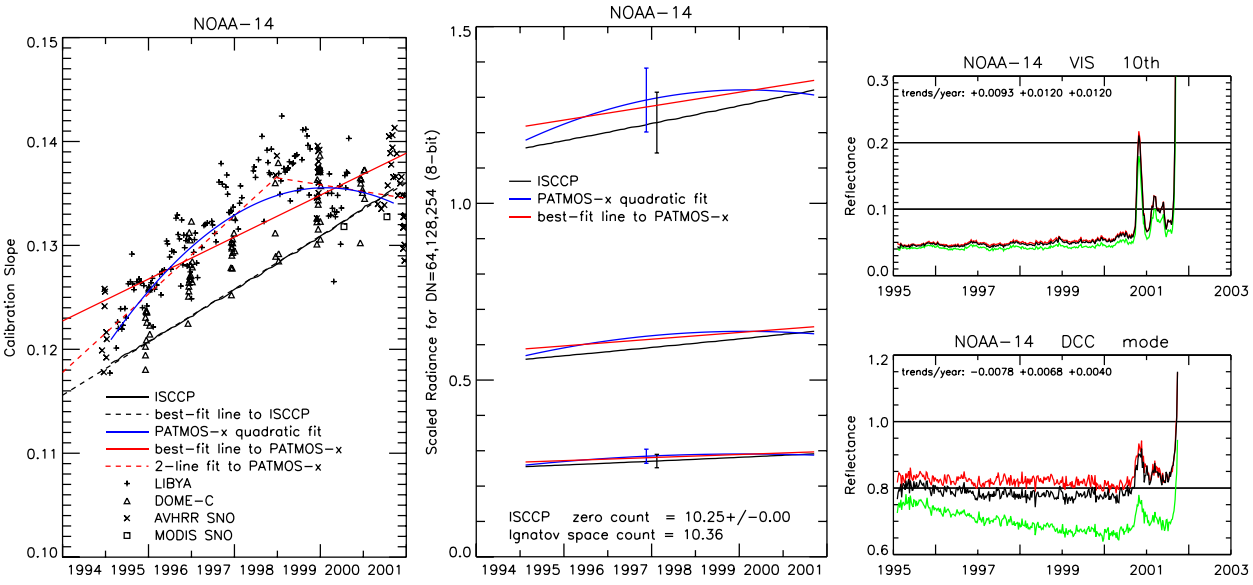


FIG. 5. (a) As in Fig. 1a, but for *NOAA-14*. The AVHRR SNO results are from *NOAA-11* (proceeding) and *NOAA-16* (following). The "□" symbols show the results from matched MODIS SNO observations. (b) As in Fig. 1b, but for *NOAA-14*. (c) As in Fig. 1c, but for *NOAA-14*.

improves the comparison), but GI is about 5% smaller than GP, translating into a scaled radiance difference of about 3% at midrange [Fig. 4b, where the average of OI is equivalent to an 8-bit space count of 9.80 as compared to 10.25 from Ignatov et al. (2005)]. Figure 4b also shows that the time variations of GI shown in Fig. 4a were partly compensated by variations in OI, producing only very small changes in the scaled radiances. The PATMOS-x quadratic fit seems driven in part by the SNO points at the beginning of the record, consistent with a suspected decrease in sensor sensitivity at the beginning. If these early points are removed, the best linear fit is very similar to the linear fit to ISCCP results with a bias that is within the uncertainties (error bars in Fig. 4b). These results and a retrospective review of all the vicarious calibration results for *NOAA-11* suggest that its time record can be represented just as well by a simpler linear fit to all of the results (Table 5). The original calibration produces an overall trend in 10th of 0.48 but only 0.11% yr^{-1} in DCC. The ISCCP calibration increases the trend at 10th slightly to 0.59% yr^{-1} but causes a decreasing trend at DCC of −0.39% yr^{-1}. Using the best fit to the PATMOS-x points produces about the same trend of +0.52% at 10th and −0.50% yr^{-1} in DCC (Fig. 4c). We tested the trend statistics shown in Fig. 4c by removing the Pinatubo period (June 1991–March 1993); this did not change the results significantly. Overall, the ISCCP and PATMOS-x calibrations for this difficult case are about the same quality and agree to within their uncertainties.

NOAA-14 is the oldest and longest in storage before launch of the five-channel AVHRRs built in the 1980s and early 1990s. Figure 5a shows that its calibration appears to exhibit very nonlinear behavior over its lifetime. The overall trend of the ISCCP time record (black solid line is the actual variation of GI, dashed black line is the best linear fit) is very similar to that obtained from a linear fit to the PATMOS-x data points (red solid line; Table 5); the blue solid curve is the PATMOS-x quadratic fit. Figure 5b shows that the ISCCP calibration, which varies linearly over the record, agrees better with the PATMOS-x result at the beginning and end of the record but is about 3% smaller at midrange in the middle of the record [OI is equivalent to an 8-bit space count of 10.25 as compared to 10.36 from Ignatov et al. (2005)]. However, this difference is still within the uncertainties of both results, shown by the error bars in Fig. 5b. Figure 5a also shows a bilinear fit to the *NOAA-14* record (red dashed line) that fits the PATMOS-x data as well (same standard deviation of about 2%) as the quadratic fit (Table 5). Again, the very last SNO results for PATMOS-x, when the orbit has drifted to the terminator (Table 4), have an important effect on the quadratic fit; in addition, the points from the Libyan desert in 1998–99 also seem different than the rest of the data. If these points are neglected, then a linear fit to the whole record is just as reasonable and has a slope even closer to the ISCCP trend (Table 5). Figure 5c shows the original, the ISCCP, and the best-bilinear-fit PATMOS-x results for 10th and DCC. The

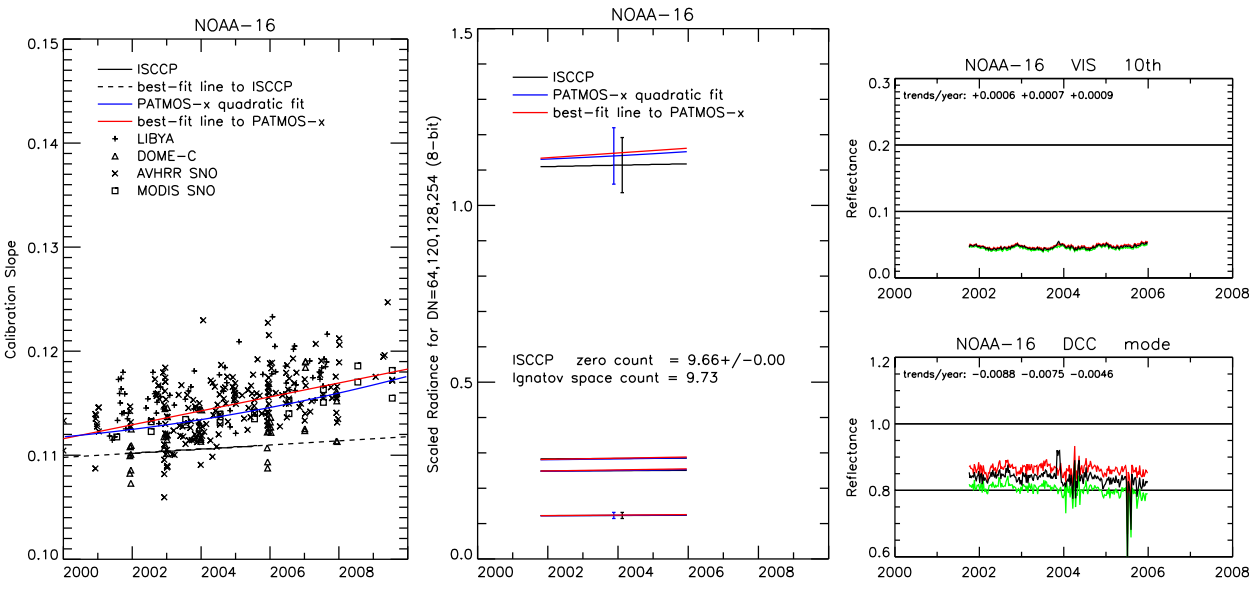


FIG. 6. (a) As in Fig. 1a, but for *NOAA-16*. The AVHRR SNO results are from *NOAA-14* (proceeding) and *NOAA-18* (following). The "□" symbols show the results from matched MODIS SNO observations. (b) As in Fig. 1b, but for *NOAA-16*. (c) As in Fig. 1c, but for *NOAA-16*.

scale in this figure does not show the nonlinear behavior well, but it does show a change in behavior during the last two years. The appearance of strong seasonal variations is caused by the rapidly increasing solar zenith angle—the *NOAA-14* orbit drifted to one of the latest times of day of all of the afternoon satellites—and the strong anisotropy of the ocean reflectances. The effect at DCC is caused by strong variations of the sample size. The orbital geometry reduced the number of pixels available with $\mu_0 \geq 0.4$ by a factor of 4 toward the end of the record, reinforcing the suggestion that the last SNO points in the PATMOS-x results might be biased. If we remove the data after August 2000 from Fig. 5c, then the original calibration gives trends at 10th and DCC of $+0.1\%\ \mathrm{yr}^{-1}$ and $-2.0\%\ \mathrm{yr}^{-1}$, respectively (compared to the very different values shown in Fig. 5c). Likewise, in contrast to the trends shown in Fig. 5c, without the later part of the record, the ISCCP calibration gives $+0.2\%\ \mathrm{yr}^{-1}$ for 10th and reduces the DCC trend magnitude to $-0.5\%\ \mathrm{yr}^{-1}$. The best bilinear fit to PATMOS-x gives $+0.5$ and $-0.1\%\ \mathrm{yr}^{-1}$ at 10th and DCC, respectively.

NOAA-16 is the other AVHRR directly calibrated by MODIS (there are also two comparisons for *NOAA-14*). Figure 6a shows that the degradation rate inferred by ISCCP (black solid line is actual and dashed black line is the best linear fit) is a little smaller than the best linear fit to the PATMOS-x data points (red solid line), which is as good a fit (2% scatter standard deviation; Table 5) and differs only slightly from the PATMOS-x quadratic fit (blue solid curve). Figure 6b shows that all of these results give the same scaled radiances over most of the range, with the ISCCP values smaller than the PATMOS-x results by only 2%–3% at the largest values [OI is equivalent to an 8-bit space count of 9.66, nearly the same as 9.73 from Ignatov et al. (2005)]. Several anomalies occurred in the operation of this AVHRR during 2003–05, but apparently they did not affect the calibration. As with *NOAA-18*, a fit to the MODIS points alone lowers the gain by 1%–2%. These results agree to within the uncertainties, as illustrated by the error bars in Fig. 6b. Figure 6c shows that the original, ISCCP, and best-fit results to PATMOS-x are all equivalent, showing almost no trend at 10th but the PATMOS-x best-fit result is slightly better for DCC (trends = -0.88, -0.75, and $-0.46\%\ \mathrm{yr}^{-1}$, respectively). If only the MODIS points are used to compare with *NOAA-16*, then the result is slightly lower than the PATMOS-x result by about 1%.

d. *Morning satellites (*NOAA-8, NOAA-10, NOAA-12, NOAA-15, *and* NOAA-17*)*

NOAA-8 only contributed a little over one year of data to the ISCCP collection (Table 1) and, given its very early morning orbit (Table 4), it was very difficult to calibrate relative to *NOAA-7* (the standard at the time). Figure 7a shows that, although the PATMOS-x analysis (blue solid curve is the quadratic fit, red solid line is the best linear fit) was applied to a much longer time record, most of its information is concentrated near the end of *NOAA-8*'s lifetime, whereas the ISCCP analysis is concentrated at the beginning (black solid line is the actual GI, and black dashed line is the best linear fit to

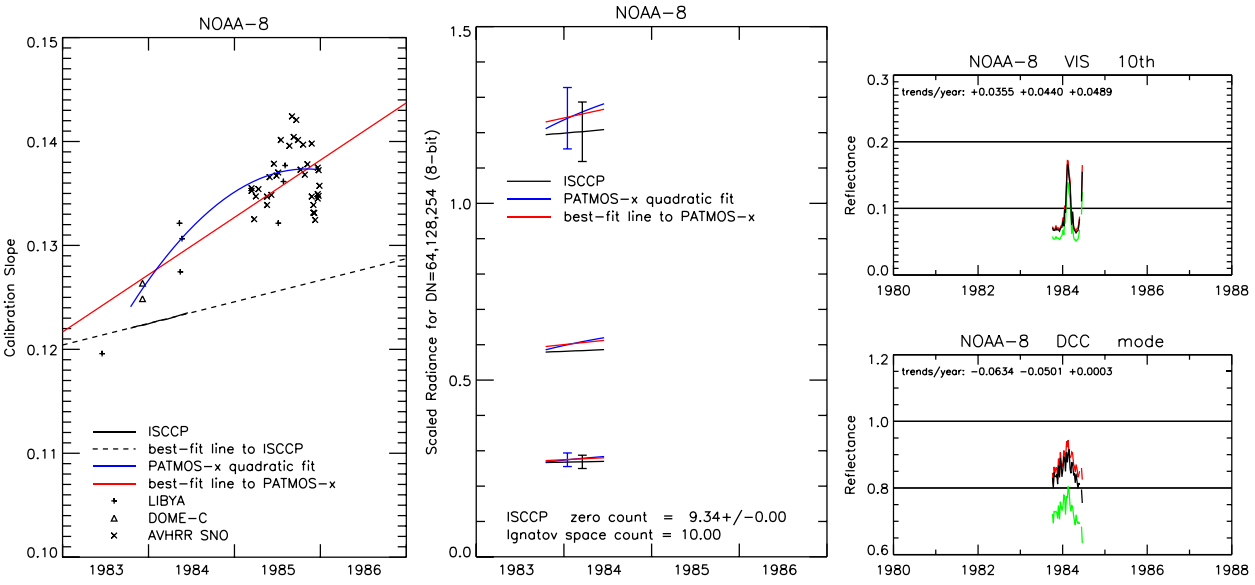


FIG. 7. (a) As in Fig. 1a, but for *NOAA-8*. The AVHRR SNO results are from *NOAA-7*. (b) As in Fig. 1b, but for *NOAA-8*. (c) As in Fig. 1c, but for *NOAA-8*.

the results in 1983–84 extrapolated to the end of 1986). This concentration of points at the end of the period contributes to the nonlinear PATMOS-x fit, but the best linear fit to the same data is nearly as good as the quadratic fit (scatter standard deviations both about 2%; Table 5). Even though the degradation rate inferred from the linear fit to the longer record is larger than determined by ISCCP, Fig. 7b shows that the scaled radiance differences over the portion of the record included in the ISCCP record are partially compensated by the OI value and are only a few percent lower at the end of this period, well within the estimated uncertainty. This shows that the procedure for calibrating *NOAA-8* relative to *NOAA-7* worked fairly well [the average OI is equivalent to an 8-bit space count of 9.34 vs 10.0 from Ignatov et al. (2005)]. Figure 7c illustrates the difficulties with the morning orbits in that reflectance statistics, even when the solar zenith angle is constrained, exhibit strong seasonality (e.g., 10th) because the geographic coverage and the number of available pixels vary dramatically over each year, the latter by a factor of 3. Even the DCC statistics exhibit this behavior. Still the DCC results show that the best linear fit to PATMOS-x eliminates the original time trend over the short record of $-6.34\%\ \mathrm{yr}^{-1}$, whereas the trend with the ISCCP calibration is still $-5.01\%\ \mathrm{yr}^{-1}$.

NOAA-10 is the oldest and longest in storage before launch of the four-channel AVHRRs built in the early 1980s. Figure 8a shows that both the ISCCP (solid black line is the actual variation of GI) and PATMOS-x (blue solid curve) results indicated a nonlinear variation of the sensor with a stronger degradation for the first two years followed by a smaller degradation or even a small increase in sensitivity for the next two years. This change in behavior occurred after the instrument underwent a degassing procedure in May 1989. The actual variation of GI agrees well with the quadratic variation of GP, except that GI shows a weak continued degradation, whereas GP implies an increased sensitivity in the latter part of the record. Figure 8a also shows a bilinear fit to the PATMOS-x points (red dashed line), which is just as good as the quadratic fit (scatter standard deviation about 2%; Table 5). However, the magnitude of the nonlinearity in these fits may be exaggerated by the single set of Dome C points in early 1989, which were obtained just before the instrument outgassing event. If these points are neglected, then the bilinear and best linear fit (red solid line) to the PATMOS-x points are nearly the same (Table 5).The overall (best linear fit to the) ISCCP degradation rate (dashed black line) is very similar to that given by the linear fit to the PATMOS-x points. Figure 8b shows that the differences in scaled radiances among the two nonlinear calibrations and the best linear fit (without the Dome C points in 1989) translate into 3%–5% differences at midrange, but unlike every case considered so far, the ISCCP calibration gain produces larger radiance values even in the lower range [the average OI is equivalent to an 8-bit space count of 8.25 as compared to 9.01 from Ignatov et al. (2005)]. Figure 8b also shows that a change of value of OI in May 1989 (Table 3) partially compensates for the lower GI introduced at the same time. Figure 8c shows the 10th and DCC variations: neither the ISCCP result nor the best bilinear fit to the PATMOS-x points

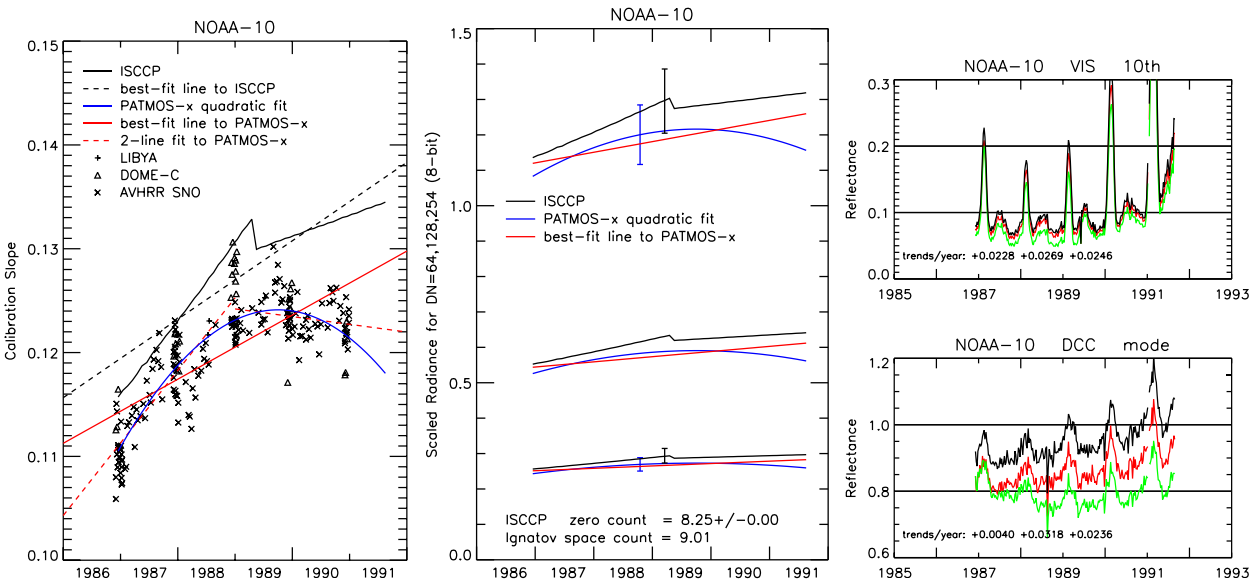


FIG. 8. (a) As in Fig. 1a, but for *NOAA-10*. The AVHRR SNO results are from *NOAA-9* and *NOAA-11*. (b) As in Fig. 1b, but for *NOAA-10*. (c) As in Fig. 1c, but for *NOAA-10*.

appears to work well, both increasing a positive trend in the original calibration from $+2.28\%\ \mathrm{yr}^{-1}$ and $+0.40\%\ \mathrm{yr}^{-1}$ for 10th and DCC, respectively, to $+2\%$–$3\%\ \mathrm{yr}^{-1}$. However, these results are suspect both because of the strong seasonal variation of the number of available targets and the drift of the orbit, which reduced the available targets with large enough solar zenith angles by almost a factor of 2 over the record. We repeated the analysis by requiring >5000 pixels per week, effectively removing the segments of the record corresponding to the peaks in 10th and DCC. Although this reduces the trend magnitudes to about $1\%\ \mathrm{yr}^{-1}$, it does not change the conclusion.

There were many problems with the *NOAA-12* AVHRR in midlife that led to the following events: 1) the operating temperature was raised on 2 January 1992 to above specification to improve scan motor performance but was returned to normal on 28 January 1992, 2) a midlife degassing was performed on 12 October 1993, and 3) the operating temperature was again raised above the specified limit for 28 July–3 August 1994. The ISCCP procedure tried to account for these changes (Table 3), based on observations over very short time periods, producing the dramatic changes in gain (and offset) illustrated in Fig. 9a (solid black line). The statistics available from the short time periods before and after these operational changes, combined with one of the most extreme orbit geometries and drift rates that eliminated convective cloud targets with $\mu_0 > 0.4$ after 1997—this satellite actually crossed the terminator at the end of its life (Table 4), produced a very poor characterization of the calibration with the ISCCP methodology. Even allowing lower μ_0 observations into the statistics (all $\mu_0 \geq 0.2$ were used) or constraining μ_0 to a narrower range ($0.2 \leq \mu_0 \leq 0.5$) did not improve the results. Nevertheless, as Fig. 9a shows, the best linear fit to the evolution of GI (dashed black line), although about 5% lower on average, is actually similar in trend to the best linear fit to the PATMOS-x results (red solid line), which fits the data and the PATMOS-x quadratic fit (blue solid curve) as Table 5 shows. Figure 9b shows that dramatic changes of GI are partially offset by changes in OI, producing only small variations of scaled radiances: the midrange differences between the ISCCP and PATMOS-x radiances at the beginning and end of the *NOAA-12* lifetime are only about 6%–8% [the average OI is equivalent to an 8-bit space count of 9.72 as compared to 10.25 from Ignatov et al. (2005)]. These differences are not much larger than the uncertainties, as shown by the error bars in Fig. 9b. The large solar zenith angles and drift of the orbit produces even stronger seasonality for the 10th percentile values over water and a strong positive trend in the statistics (Fig. 9c)—even with the strong degradation suggested by the PATMOS-x results (Fig. 9a). The trend with the original calibration is $+3.8\%\ \mathrm{yr}^{-1}$ at 10th, which the ISCCP and PATMOS-x correction make even worse $>+5\%\ \mathrm{yr}^{-1}$. The DCC results seem more stable but still show a seasonality not seen for any other satellite: the number of DCC targets falls essentially to zero by 1997. For DCC, the trend statistics for the original calibration is $+0.1\%\ \mathrm{yr}^{-1}$, which the ISCCP and best fit to PATMOS-x corrections increase to about $+2\%$–$3\%\ \mathrm{yr}^{-1}$. If we neglect the data after 1997, then the trends with original,

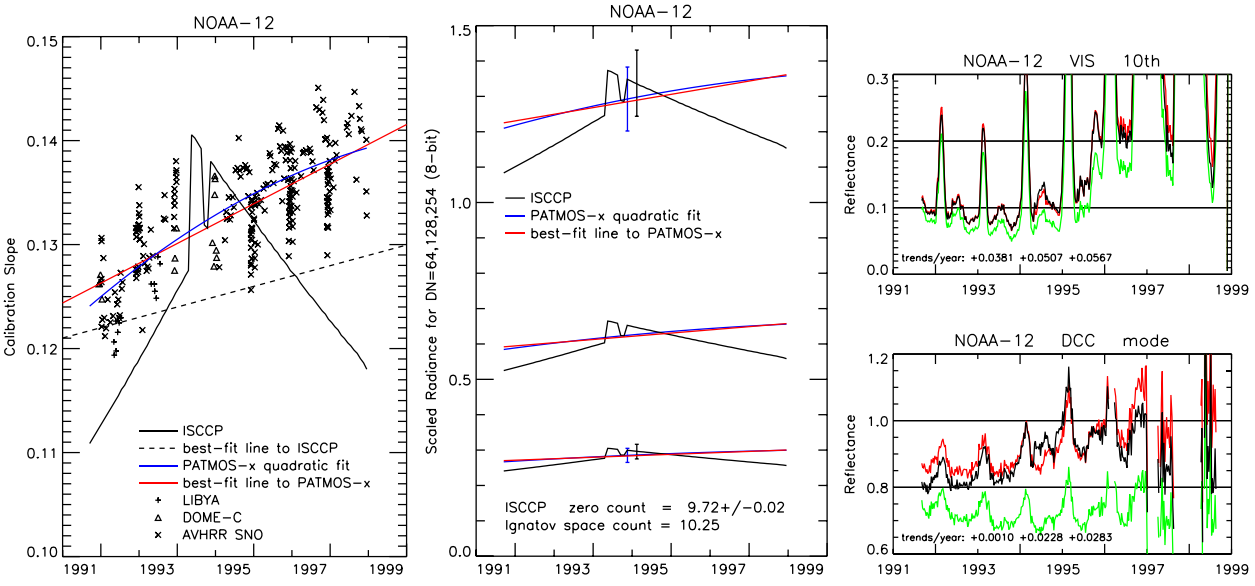


FIG. 9. (a) As in Fig. 1a, but for *NOAA-12*. The AVHRR SNO results are from *NOAA-11* and *NOAA-14*. (b) As in Fig. 1b, but for *NOAA-12*. (c) As in Fig. 1c, but for *NOAA-12*.

ISCCP, and PATMOS-x fit are again reduced to about 1% yr^{-1}. Clearly, the changing and extreme solar geometry makes it appear as if there is actually an increase in sensor sensitivity.

Figure 10a compares the ISCCP and PATMOS-x results for *NOAA-15*, which is another very short record in the ISCCP collection. This is the first of the AVHRRs with a bilinear visible channel response function. Because of the significantly larger gain in the ISCCP calibration (almost 20% during the period used) and an OI equivalent to nearly the same (8 bits) space count as given by Ignatov et al. (2005) (9.63 vs 9.50), the scaled radiances are 5% larger at midrange (Fig. 10b), though still within estimated uncertainties. The ISCCP curves shown in these figures were generated directly from the calibration coefficients that were supposed to be used to generate the calibration table used in the cloud analysis; however, a review of all tables revealed that the table for *NOAA-15* was produced using only the high-range gain and not the bilinear gain. This had the effect of reducing

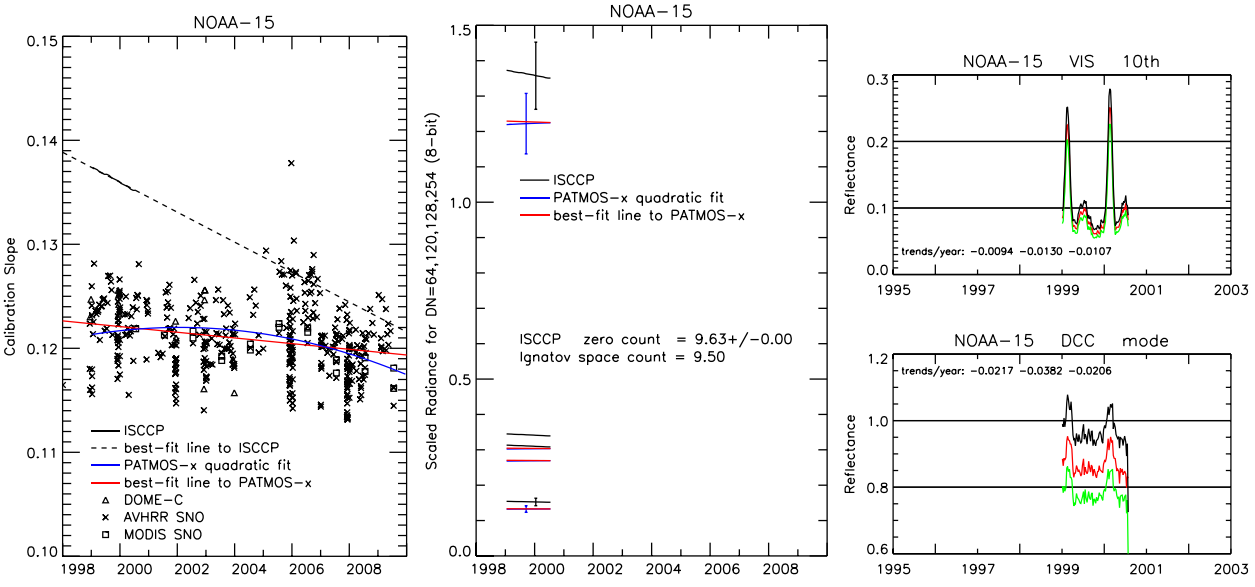


FIG. 10. (a) As in Fig. 1a, but for *NOAA-15*. The AVHRR SNO results are from *NOAA-14*, *NOAA-16* and *NOAA-18*. (b) As in Fig. 1b, but for *NOAA-15*. (c) As in Fig. 1c, but for *NOAA-15*.

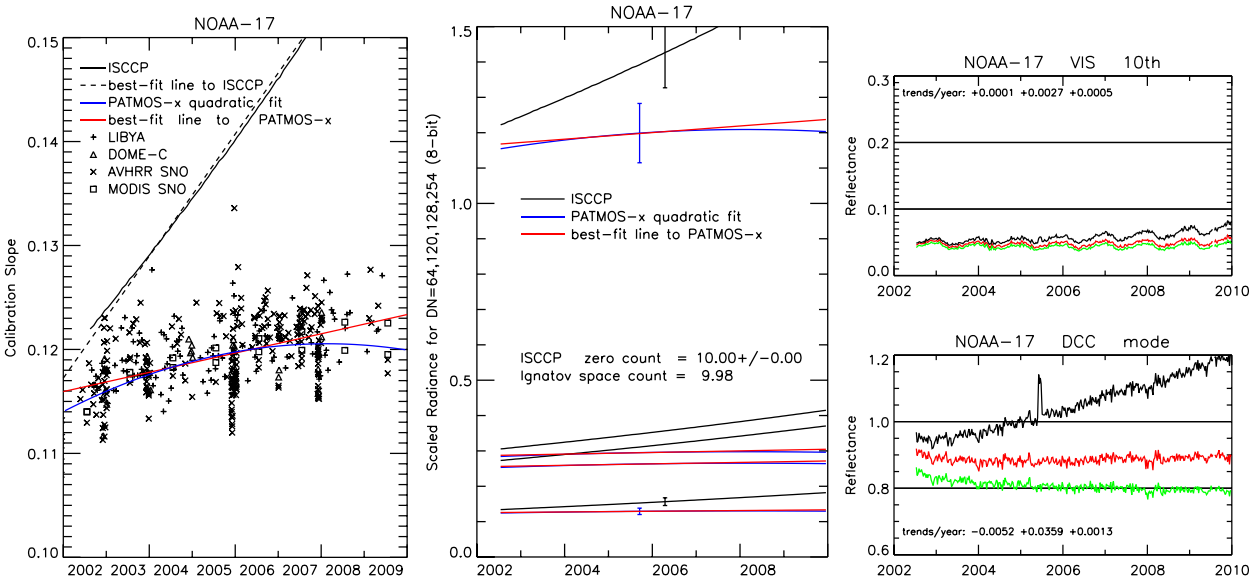


FIG. 11. (a) As in Fig. 1a, but for *NOAA-17*. The AVHRR SNO results are from *NOAA-16* and *NOAA-18*. The "□" symbols show the results from matched MODIS SNO observations. (b) As in Fig. 1b, but for *NOAA-17*. (c) As in Fig. 1c, but for *NOAA-17*.

all lower radiances to zero, but since the morning satellite is rarely used to produce the ISCCP gridded products, the resulting error in the cloud results is small. Although the calibration trend is not accurate and despite using only the higher gain in the nominal calibration, the final calibration for the period of use by ISCCP actually agrees with PATMOS-x within the uncertainties (Fig. 10b). The PATMOS-x results again benefit from a much longer time record for determining the evolution of the gain (blue solid curve in Fig. 10a) and the best linear fit (red solid line) is equally good (Table 5). Figure 10c shows the usual problems associated with the early morning orbit geometry; the ISCCP calibration increases the trend in DCC from about 2% yr^{-1} to about 4% yr^{-1}, but the best fit to PATMOS-x produces little improvement in the apparent trend.

Figure 11a shows that the ISCCP estimate of sensor evolution for *NOAA-17* is very different from the PATMOS-x result, leading to midrange scaled radiance differences of about 10% at the end of the record (Fig. 11b), even though the average OI is equivalent to an 8-bit space count of 10.00 compared with 9.98 reported by Ignatov et al. (2005). This satellite is unusual among the morning satellites because it was launched into an orbit much later in the morning (Table 4), 1000 LT instead of 0700–0800 LT, which suppresses the seasonal variations seen in the 10th statistics shown in Fig. 11c. In DCC, the original NOAA calibration indicates a degradation of −0.52% yr^{-1}, which the best linear fit to PATMOS-x (red solid line in Fig. 11a, which is as good as the quadratic fit, blue solid curve in Fig. 11a; Table 5) reduces to +0.1% yr^{-1}. The ISCCP result, however, caused a +3.59% yr^{-1} trend for *NOAA-17*. Investigating all the factors that caused problems with calibrating morning satellites, we found that this result is actually based on only the first few years of data and extrapolated to the end of 2009, not on a fit to the whole record as done for other satellites, producing a much less accurate estimate of the calibration trend (like *NOAA-8* and *NOAA-15*).

5. Summary

The calibrations from ISCCP and PATMOS-x generally agree to within their estimated uncertainties (about ±7% for radiances), except for *NOAA-17* (*NOAA-10* and *NOAA-15* differences are just barely larger than the sum of the uncertainties). Notably, in all cases of good agreement, the ISCCP gain values are slightly smaller than the PATMOS-x values (generally the afternoon satellites) and in the two cases of larger disagreement (and *NOAA-10*), all morning satellites, the ISCCP gain values are larger than the PATMOS-x values (but not *NOAA-8* or *NOAA-12*). The proper interpretation of "agree to within their estimated uncertainties" is that we cannot choose which result is superior given the very different approaches.

The fact that the ISCCP gain values are generally smaller than those from PATMOS-x may be explained by the emphasis of the former on darker targets and the latter on brighter targets (since the offsets are nearly the same). Moreover, the average mode values of the reflectance distributions from the DCC analysis produced by the ISCCP calibration are generally smaller than the values produced by the PATMOS-x calibration by a few

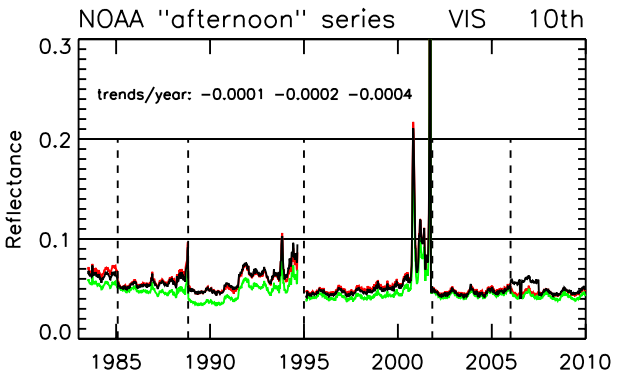
TABLE 6. Average percent change and the time derivative (% yr^{-1}) of the change of the gain values from the original ISCCP values, as well as the initial in time values of the gain and space count CTsp in 8-bit counts and the time derivative of the gain (% yr^{-1}) for the new calibration. The initial gain values for the bilinear AVHRRs (*NOAA-15* through *NOAA-18*) are for the lower range. In addition, an offset correction subtracting 0.01 was made for *NOAA-18* from January 2006 through June 2007.

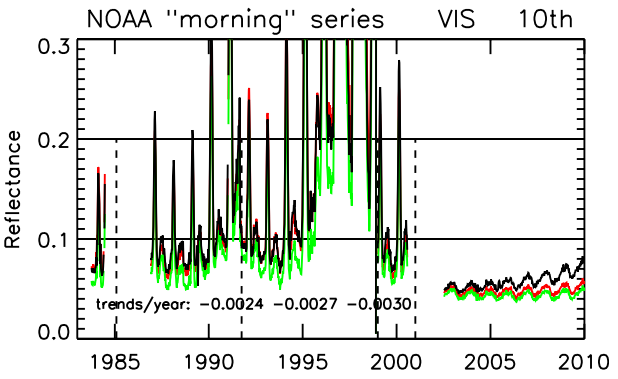
Satellite	Average gain change (%)	Gain change derivative (% yr^{-1})	Initial gain	Initial CTsp	Derivative (% yr^{-1})
NOAA-9	+1.686	+0.240 326	0.004 283 74	9.00	+6.811 62
NOAA-18	+1.473	−0.758 436	0.004 426 90	9.80	+1.945 82
NOAA-7	+1.254	+1.211 53	0.004 646 38	7.79	+2.527 06
NOAA-11	+2.257	−1.676 59	0.004 452 16	9.81	+0.364 933
NOAA-14	+2.001	−0.281 943	0.004 885 44	10.25	+1.840 79
NOAA-16	+1.666	+0.186 576	0.004 463 35	9.66	+0.347 305
NOAA-8	+2.105	+1.430 02	0.004 961 39	9.34	+3.240 32
NOAA-10	−3.663	−0.459 976	0.004 618 68	8.25	+2.945 29
NOAA-12	+3.120	+0.289 517	0.004 944 10	9.72	+1.164 66
NOAA-15	−10.362	+0.722 400	0.004 898 10	9.63	−0.269 631
NOAA-17	−15.380	−2.829 71	0.004 666 95	10.00	+0.797 033

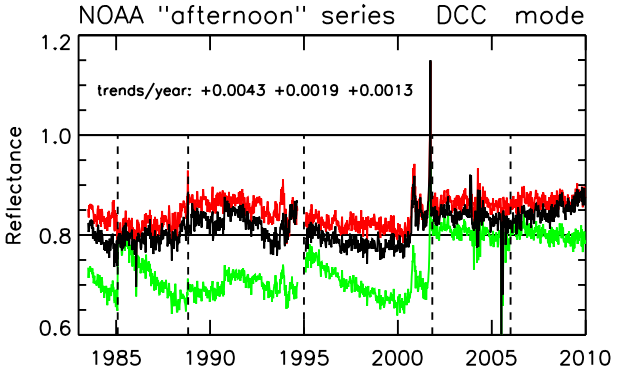
percent. We note, however, that the PATMOS-x gains are slightly larger than the MODIS-only gains. The ISCCP procedure uses a global distribution of 28 geographic regions, each divided by surface or vegetation type (including ocean), as targets composed of millions of samples (Brest and Rossow 1992; Brest et al. 1997), which tends to weight the smaller radiances more. On the other hand, the PATMOS-x procedure, with many fewer samples (see figures), emphasizes the larger radiances by using a desert and ice sheet as targets (Heidinger et al. 2010 or Molling et al. 2010). The ISCCP results use top-of-atmosphere reflectances with no adjustments for differences in instrument spectral responses, except the differing solar constants used to determined scaled radiances. The PATMOS-x results employ a radiative transfer model to correct for spectral differences and atmospheric effects.

Given that 1) the average differences between the results are smaller than the estimated uncertainties in either result (suggesting the spectral and target anisotropy effects are small), 2) the analysis methods are very different and emphasize different ranges of radiances, and 3) the records are best anchored near the beginning for ISCCP and near the end for PATMOS-x, we argue that the best use of this complementary information is to average the two results. One difference from the PATMOS-x results is that we use the best linear fit to the points shown in the figures, as we could not establish in most cases that there is any statistical difference between a linear and quadratic fit. In two cases, *NOAA-10* and *NOAA-14*, there is some evidence for nonlinear behavior; but the deviations from a ''good'' linear fit are well within the estimated uncertainties, and there are some reasons to suspect the points that seem to influence the nonlinear fits most strongly. In two cases, *NOAA-15* and *NOAA-17*, the ISCCP results are clearly incorrect, so we adopt the PATMOS-x results (best linear fits) for these two satellites. Accordingly, the quantities in (5), namely, δ_0 and $\partial G/\partial t$, are determined for the average of the ISCCP and PATMOS-x gains for each month except for *NOAA-15* and *NOAA-17*. A summary of the proposed calibration is given in Table 6.

Figure 12 summarizes the calibration for the whole series of afternoon AVHRRs: there is little difference among the calibrations at the lower reflectances (and even at larger percentiles in the midrange), but the differences in DCC are more readily apparent. The original ISCCP calibration, which was based on land (and ocean) reflectance statistics concentrated in the lower range of scaled radiance values, successfully removed most of the sensor degradation for individual AVHRRs. The overall linear trend was reduced from +0.43% yr^{-1} with the original NOAA calibration (equivalent to about a 12% increase in reflectance over 27 years—but note that each satellite exhibits larger sensor degradations over its lifetime, so it is only the generally larger base line for *NOAA-16* and *NOAA-18* relative to *NOAA-7* and *NOAA-9* that results in a positive trend) to +0.19% yr^{-1} with the original ISCCP calibration (equivalent to a 5% increase over 27 years). The best linear fit to PATMOS-x reduces the trend in DCC to 0.13% yr^{-1} (3.5% increase over 27 years). Using the proposed calibration changes discussed in section 4 and above, the overall trend is further reduced to +0.1% yr^{-1}, equivalent to absolute reflectance decrease <3% over 27 years. The larger values occurring in the middle part of the *NOAA-11* period are contaminated by Mount Pinatubo aerosols; this portion of the results is excluded from the trend calculation. Likewise, the sudden increase near the end of the *NOAA-14* lifetime is associated with a dramatic decrease in the available sunlit pixel number as the orbit drifts toward the terminator; these points are also excluded. Note also that *NOAA-14*, which exhibits evidence of nonlinear calibration changes, is the afternoon







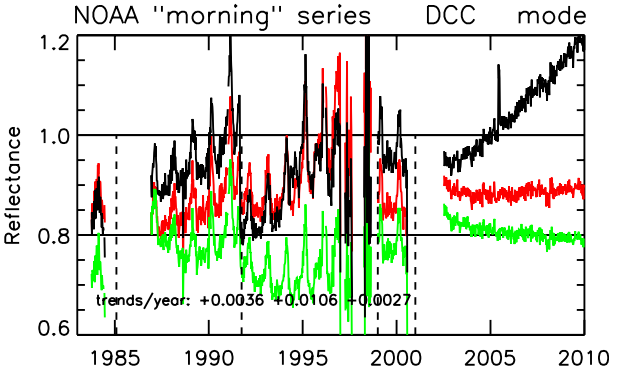


FIG. 12. Comparison of (top) the time variations of 10th percentile visible reflectance values from observations collected over the global oceans for each week with cosine solar zenith angles μ_0 constrained to be $\geq$0.4 and (bottom) the time variations of the mode of the distribution of visible reflectances for DCC identified by infrared brightness temperatures $<$210 K. The green lines show the results for the original calibration for all of the afternoon AVHRRs (*NOAA-7*, *NOAA-9*, *NOAA-11*, *NOAA-14*, *NOAA-16*, and *NOAA-18*), the red lines show the results with the best linear fit to the PATMOS-x calibration points in Fig. 1a, and the black lines show the results with the original ISCCP calibration. The vertical dashed lines indicate changes of satellite. The linear reflectance trend per year for the three calibrations (original, ISCCP, PATMOS-x) is indicated.

FIG. 13. Comparison of (top) the time variations of 10th percentile visible reflectance values from observations collected over the global oceans for each week with cosine solar zenith angles μ_0 constrained to be $\geq$ 0.4 and (bottom) the time variations of the mode of the distribution of visible reflectances for DCC identified by infrared brightness temperatures $<$ 210 K. The green lines show the results for the original calibration for all of the morning AVHRRs (*NOAA-8*, *NOAA-10*, *NOAA-12*, *NOAA-15*, and *NOAA-17*), the red lines show the results with the best linear fit to the PATMOS-x calibration points in Fig. 1a, and the black lines show the results with the original ISCCP calibration. The vertical dashed lines indicate changes of satellite. The linear reflectance trend per year for the three calibrations (original, ISCCP, PATMOS-x) is indicated.

AVHRR that was longest in storage before flight. The slightly larger values for *NOAA-16* and *-18* relative to the baseline of the earlier satellites are associated with their orbits being closer to local noontime.

Figure 13 summarizes the whole morning AVHRR record, where the situation is less clear. Despite the seasonal variations in 10th, associated with the more extreme solar geometry, the overall trend is the same in the original calibration, for the ISCCP calibration and for the revised ISCCP calibration, all about $-0.3\%\ \mathrm{yr}^{-1}$. For DCC, the original calibration produces an overall trend of only $+0.36\%\ \mathrm{yr}^{-1}$, which the original ISCCP calibration increased to $+1.06\%\ \mathrm{yr}^{-1}$, mostly because of the *NOAA-15* and *NOAA-17* results. The revisions proposed and discussed in section 4 and above return the overall trend to a little less than $+0.3\%\ \mathrm{yr}^{-1}$ (equivalent to about 7% over 27 years); however, this result appears to be produced mostly by orbit drift changes in solar illumination. Note also that *NOAA-10*, which exhibits evidence of nonlinear calibration changes, is the morning AVHRR that was longest in storage before flight.

Table 6 indicates the average magnitude of the percent changes to the gain and its time derivative and the initial values of the new average gain and space count (for 8-bit counts).

6. Discussion

Although the ISCCP and PATMOS-x methods of propagating the anchoring calibration to the rest of the

AVHRR series seem different, they both rely on somewhat similar assumptions about the behavior of the instruments and the whole earth or some part of it as reflectance targets. Both assume that there are no shorter-term variations in the calibration, so that a statistical treatment of the radiance/reflectance statistics over the lifetime of each instrument is proper. The ISCCP method obtains monthly results from a very large number of targets comprising the whole earth and assumes that their aggregate behavior is constant over the whole record, whereas the PATMOS-x method obtains results about every six months from two small targets that are also assumed to be constant in time. The ISCCP method attempted to fit the variations on time scales shorter than the 3–6-yr lifetimes of the satellites, whereas the PATMOS-x method assumes a single fit to each lifetime time series. None of these assumptions was known to be true a priori, but the excellent quantitative agreement of the two results provides a posteriori confirmation of the constancy of the target reflectances and the slow, generally steady evolution of instrument sensitivity. In some cases, where the ISCCP results implied changes in behavior (*NOAA-11* and *NOAA-12*), this retrospective comparison shows that a simpler evolution is an equally good fit to the data. In some cases (*NOAA-10* and *NOAA-14*), there is a suggestion of changing behavior,where the changes correspond to some documented events. This behavior is represented by a strong quadratic fit in the PATMOS-x results or could equally well be represented by a bilinear fit as we have shown. The scatter of the measurements around these fits is interpreted by both analyses as measurement error (method uncertainty) but could be indicative of shorter-term variations of the instrument. Given the magnitude of the scatter (uncertainty, about 2%–3%), a linear fit to all of the results, even for *NOAA-10* and *NOAA-14*, seems sufficient. A more complicated representation may be confusing method uncertainty for real variations. For this reason we also argue that the best calibration is the average of the two results except for the two cases where there is strong disagreement (*NOAA-15* and *NOAA-17*) and where the PATMOS-x results correspond better to the independent check using the DCC method.

Given the interpretation of these two analyses, we propose answers to the following questions:

1) How much of the observed variation is methodological noise and how much real calibration variation on shorter time scales? On time scales less than the lifetime of each instrument (3–6 years), the scatter of results compiled on time scales greater than one month is about 2%–3% in radiance, which seems to be methodological ''noise'' or measurement ''error.''
2) How frequently must the radiometer performance be measured? Although the evolution of the instrument sensitivity generally seems to be steady and slow, this should not be assumed to be true. The noise level in the determinations of the calibration also suggests that monthly to subyearly measurements are required for a statistically robust result. More data are better.
3) What is a sufficient sample size for monitoring small changes? Checking the calibration every month appears to be sufficient to suppress the noise, but the accuracy of determining the degradation rate appears to depend on the length of record used to estimate it: to attain accuracies of at least 0.1% yr^{-1} requires more than a few years of monthly results.

The comparison of these two results provides mutual confirmation of their records in that they generally agree to within their estimated uncertainties. Given this interpretation, the averaged results provide an absolute calibration of the visible channels to within about ±3% in radiance (the PATMOS-x analysis is the only one that also provides results for the near-IR channel). The PATMOS-x analysis, unlike ISCCP, provides results for the AVHRRs on TIROS-N, *NOAA-6* at the beginning of the record, and for *NOAA-19* and *MetOp-A* most recently. Planned revisions of the ISCCP cloud products include extending the calibrations to these other instruments as well.

Acknowledgments. We thank Andrew Heidinger (NOAA CIMSS) for providing all of his data used to produce the PATMOS-x calibrations, which made possible this very detailed comparison. We also thank David Doelling (NASA Langley) for the discussions and for showing us his direct MODIS to geostationary calibration results. We thank Jack Xiong for helping us understand the MODIS calibration procedure, and Laura Hinkelman and Tom Stone for the useful discussions about these results. This work was supported by NASA Grant NNX08AJ80G (Laura Hinkelman, PI; L. Tsaoussi, PM) and by NASA Grant NNX08AL79A (MEASURES, M. Maiden, PM). This paper is dedicated to our colleague, Christopher L. (Brest) Bishop, whose sudden death in 2008 ended a long and successful analysis of the AVHRR visible calibration. The quality of his work is shown by the comparison and evaluation results presented here.

REFERENCES

Arking, A., 1964: Latitudinal distribution of cloud cover from Tiros III photographs. *Science,* **143,** 569–572, doi:10.1126/science.143.3606.569.

Brest, C. L., and W. B. Rossow, 1992: Radiometric calibration and monitoring of NOAA AVHRR data for ISCCP. *Int. J. Remote Sens.,* **13,** 235–273, doi:10.1080/01431169208904037.

——, ——, and M. D. Roiter, 1997: Update of radiance calibrations for ISCCP. *J. Atmos. Oceanic Technol.,* **14,** 1091–1109, doi:10.1175/1520-0426(1997)014<1091:UORCFI>2.0.CO;2.

Che, N., and J. C. Price, 1992: Survey of radiometric calibration results and methods for visible and near infrared channels of NOAA-7, -9 and -11 AVHRRs. *Remote Sens. Environ.,* **41,** 19–27, doi:10.1016/0034-4257(92)90057-Q.

Desormeaux, Y., W. B. Rossow, C. L. Brest, and G. G. Campbell, 1993: Normalization and calibration of geostationary satellite radiances for ISCCP. *J. Atmos. Oceanic Technol.,* **10,** 304–325, doi:10.1175/1520-0426(1993)010<0304:NACOGS>2.0.CO;2.

Doelling, D. R., L. Nguyen, and P. Minnis, 2004: On the use of deep convective clouds to calibrate AVHRR data. *Earth Observing Systems IX,* W. L. Barnes and J. J. Butler, Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 5542), 281–289, doi:10.1117/12.560047.

Frouin, R., and C. Gautier, 1987: Calibration of NOAA-7 AVHRR, GOES-5, and GOES-6 VISSR/VAS solar channels. *Remote Sens. Environ.,* **22,** 73–102, doi:10.1016/0034-4257(87)90028-9.

Fu, R., A. D. Del Genio, and W. B. Rossow, 1990: Behavior of deep convective clouds in the tropical Pacific deduced from ISCCP radiance data. *J. Climate,* **3,** 1129–1152, doi:10.1175/1520-0442(1990)003<1129:BODCCI>2.0.CO;2.

Ham, S.-H., and B. J. Sohn, 2010: Assessment of the calibration performance of satellite visible channels using cloud targets: Application to Meteosat-8/9 and MTSAT-1R. *Atmos. Chem. Phys. Discuss.,* **10,** 12 629–12 664, doi:10.5194/acpd-10-12629-2010.

Heidinger, A. K., W. C. Straka, C. C. Molling, J. T. Sullivan, and X. Wu, 2010: Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record. *Int. J. Remote Sens.,* **31,** 6493–6517, doi:10.1080/01431161.2010.496472.

Hovis, W. A., J. S. Knoll, and G. R. Smith, 1985: Aircraft measurements for calibration of an orbiting spacecraft sensor. *Appl. Opt.,* **24,** 407–410, doi:10.1364/AO.24.000407.

Hu, Y., B. A. Wielicki, P. Yang, P. W. Stackhouse, B. Lin, and D. F. Young, 2004: Application of deep convective cloud albedo observation to satellite-based study of the terrestrial atmosphere: Monitoring the stability of spaceborne measurements and assessing absorption anomaly. *IEEE Trans. Geosci. Remote Sens.,* **42,** 2594–2599, doi:10.1109/TGRS.2004.834765.

Ignatov, A., C. Cao, J. Sullivan, R. Levin, X. Wu, and R. Galvin, 2005: The usefulness of in-flight measurements of space count to improve calibration of the AVHRR solar reflectance bands. *J. Atmos. Oceanic Technol.,* **22,** 180–200, doi:10.1175/JTECH-1691.1.

Kriebel, K. T., 1981: Calibration of the METEOSAT-VIS-channel by airborne measurements. *Appl. Opt.,* **20,** 11–12, doi:10.1364/AO.20.000011.

Li, C., Y. Xue, Q. Liu, J. Guang, X. He, J. Zhang, T. Wang, and X. Liu, 2014: Post calibration of channels 1 and 2 of long-term AVHRR data record based on SeaWiFS data and pseudo-invariant targets. *Remote Sens. Environ.,* **150,** 104–119, doi:10.1016/j.rse.2014.04.020.

Matthews, E., and W. B. Rossow, 1987: Regional and seasonal variations of surface reflectance at 0.6 μm. *J. Climate Appl. Meteor.,* **26,** 170–202, doi:10.1175/1520-0450(1987)026<0170:RASVOS>2.0.CO;2.

Molling, C. C., A. K. Heidinger, W. C. Straka, and X. Wu, 2010: Calibrations for AVHRR channels 1 and 2: Review and path towards consensus. *Int. J. Remote Sens.,* **31,** 6519–6540, doi:10.1080/01431161.2010.496473.

Nagaraja Rao, C. R., and J. Chen, 1995: Inter-satellite calibration linkages for the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-7, -9, and -11 spacecraft. *Int. J. Remote Sens.,* **16,** 1931–1942, doi:10.1080/01431169508954530.

——, ——, W. F. Staylor, P. Abel, Y. J. Kaufman, E. Vermote, W. B. Rossow, and C. Brest, 1993: Degradation of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-9 spacecraft: Assessment and recommendations for corrections. NOAA Tech. Rep. NESDIS 70, 25 pp. [Available online at http://docs.lib.noaa.gov/noaa_documents/NESDIS/TR_NESDIS/TR_NESDIS_70.pdf.]

Neckel, H., and D. Labs, 1984: The solar radiation between 3300 and 12500 Å. *Sol. Phys.,* **90,** 205–258, doi:10.1007/BF00173953.

Price, J. C., 1987: Radiometric calibration of satellite sensors in the visible and near infrared: History and outlook. *Remote Sens. Environ.,* **22,** 3–9, doi:10.1016/0034-4257(87)90025-3.

Reynolds, D. W., and T. H. Vonder Haar, 1977: A bispectral method for cloud parameter determination. *Mon. Wea. Rev.,* **105,** 446–457, doi:10.1175/1520-0493(1977)105<0446:ABMFCP>2.0.CO;2.

Rossow, W. B., and A. A. Lacis, 1990: Global, seasonal cloud variations from satellite radiance measurements. Part II: Cloud properties and radiative effects. *J. Climate,* **3,** 1204–1253, doi:10.1175/1520-0442(1990)003<1204:GSCVFS>2.0.CO;2.

Slater, P. N., S. F. Biggar, R. G. Holm, R. D. Jackson, Y. Mao, M. S. Moran, J. M. Palmer, and B. Yuan, 1987: Reflectance and radiance-based methods for the in-flight absolute calibration of multispectral sensors. *Remote Sens. Environ.,* **22,** 11–38, doi:10.1016/0034-4257(87)90026-5.

Sohn, B.-J., S.-H. Ham, and P. Yang, 2009: Possibility of the visible-channel calibration using deep convective clouds overshooting the TTL. *J. Appl. Meteor. Climatol.,* **48,** 2271–2283, doi:10.1175/2009JAMC2197.1.

Staylor, W. F., 1990: Degradation rates of the AVHRR visible channel for the NOAA 6, 7 and 9 spacecraft. *J. Atmos. Oceanic Technol.,* **7,** 411–423, doi:10.1175/1520-0426(1990)007<0411:DROTAV>2.0.CO;2.

Thuillier, G., M. Hersé, D. Labs, T. Foujols, W. Peetermans, D. Gillotay, P. C. Simon, and H. Mandel, 2003: The solar spectral irradiance from 200 to 2400 *nm* as measured by the SOLSPEC spectrometer from the Atlas and Eureca missions. *Sol. Phys.,* **214,** 1–22, doi:10.1023/A:1024048429145.

Vermote, E., and Y. J. Kaufman, 1995: Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views. *Int. J. Remote Sens.,* **16,** 2317–2340, doi:10.1080/01431169508954561.

Verstraete, M. M., B. Pinty, and R. F. Dickinson, 1990: A physical model of the bidirectional reflectance of vegetation canopies 1. Theory. *J. Geophys. Res.,* **95,** 11 755–11 765, doi:10.1029/JD095iD08p11755.

Whitlock, C. H., and Coauthors, 1990: AVHRR and VISSR satellite instrument calibration results for both cirrus and marine stratocumulus IFO periods. FIRE science results 1988: Proceedings of a conference held in Vail, Colorado, NASA Conf. Publ. 3083, 141–145.

Wu, A., X. Xiong, D. R. Doelling, D. Morstad, A. Angal, and R. Bhatt, 2013: Characterization of Terra and Aqua MODIS VIS, NIR, and SWIR spectral bands' calibration stability. *IEEE Trans. Geosci. Remote Sens.,* **51,** 4330–4338, doi:10.1109/TGRS.2012.2226588.

Xiong, X., and W. Barnes, 2006: An overview of MODIS radiometric calibration and characterization. *Adv. Atmos. Sci.,* **23,** 69–79, doi:10.1007/s00376-006-0008-3.

——, J. Sun, W. L. Barnes, and V. V. Salomonson, 2007: Multiyear on-orbit calibration and performance of Terra MODIS reflective solar bands. *IEEE Trans. Geosci. Remote Sens.,* **45,** 879–889, doi:10.1109/TGRS.2006.890567.

——, ——, X. Xie, W. L. Barnes, and V. V. Salomonson, 2010: On-orbit calibration and performance of Aqua MODIS reflective solar bands. *IEEE Trans. Geosci. Remote Sens.,* **48,** 535–546, doi:10.1109/TGRS.2009.2024307.

Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, 2004: Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res.,* **109,** D19105, doi:10.1029/2003JD004457.

Zangvil, A., 1975: Temporal and spatial behavior of large-scale disturbances in tropical cloudiness deduced from satellite brightness data. *Mon. Wea. Rev.,* **103,** 904–920, doi:10.1175/1520-0493(1975)103<0904:TASBOL>2.0.CO;2.