Quart. J. R. Met. Soc. (1984), **110**, pp. 613–631 551.507.362.2:551.521.324:551.526.6

Satellite multichannel infrared measurements of sea surface temperature of the N.E. Atlantic Ocean using AVHRR/2

By D. T. LLEWELLYN-JONES, P. J. MINNETT, R. W. SAUNDERS* and A. M. ZAVODY
Rutherford Appleton Laboratory, Chilton

(Received 28 September 1983; revised 1 March 1984)

SUMMARY

The accuracy with which sea surface temperature (s.s.t.) can be measured using satellite infrared radiometers is limited primarily by uncertainties in the correction for atmospheric effects upon the measured radiance. This paper reports an investigation of the accuracy with which this correction can be made over the north-eastern and tropical Atlantic Ocean using data at 3·7, 11 and 12 μm wavelengths from the Advanced Very High Resolution Radiometer (AVHRR/2) on the NOAA-7 satellite, and the results are used as a test of atmospheric transmittance models. Simulations of atmospheric transmittances based on line-by-line calculations, using published line listings and experimental data on the water vapour 'continuum' absorption, provide regression relationships which permit the s.s.t. to be calculated from the brightness temperatures measured in each channel. New algorithms for both the 'split window' (11 and 12 μm) and the 'triple window' (3·7, 11 and 12 μm) have been derived for a range of airmasses from 1 to 2 to enable the s.s.t. to be retrieved from a 2280 km wide swath centred on the sub-satellite track. To test the validity of the simulations, the s.s.t. values derived from the satellite measurements were compared with near-coincident *in situ* measurements from oceanographic research ships. The absence of significant bias (~0·1 K) in the comparison constitutes an important experimental verification of the atmospheric transmittance model used in the simulations, and the r.m.s. difference between ship and satellite values indicates that daytime measurement of s.s.t. using the 'split-window' can be made to an accuracy of about ±0·6 K, in the eastern North Atlantic. Some possible sources of this uncertainty are discussed.

1. INTRODUCTION

The measurement of sea surface temperature (s.s.t.) from space can be regarded as a particularly demanding application of atmospheric physics. Measurements of global s.s.t. with uncertainties less than 0·5 K are required for the effective incorporation of s.s.t. data into large-scale models of global air–sea heat exchange and heat transport processes (World Climate Research Programme 1981); and in order to make such measurements from space it is necessary to use models for the processes of absorption, emission and scattering of radiation by the atmosphere which inevitably affect the radiometric measurement. The precise measurements of s.s.t. from space is a critical experimental test of such models.

In the thermal infrared (λ = 3–14 μm) the emissivity of sea water is high and relatively constant, instrumentation is well developed, and radiometric efficiency is particularly high on account of the nature of the Planck function at temperatures near 300 K. In these circumstances by far the largest source of measurement uncertainty remains in the estimation of the atmospheric correction — usually expressed as a temperature deficit ranging from under 1 K in high latitudes to over 20 K in the tropics. For this reason it is essential to use valid models of atmospheric transmittance in retrieving s.s.t. from space-borne radiometric measurements.

Since the advent of operational space-borne multichannel infrared radiometers the possibility has existed of estimating the atmospheric correction from a comparison of measured brightness temperatures in two or more wavebands, notably 3·7 and 11 μm in the case of the Advanced Very High Resolution Radiometer (AVHRR) on the TIROS-N and NOAA-6 satellites, and of 3·7, 11 and 12 μm in the case of AVHRR/2 on NOAA-7 (launched June 1981). Such measurements from AVHRR/2 are the basis of the m.c.s.s.t. (multi-channel sea surface temperature) charts of global s.s.t. issued routinely by NOAA/NESDIS. The wavelengths are chosen so that atmospheric effects

* Present address: Meteorological Office, Bracknell

are different for each channel (Prabhakara *et al.* 1974; McMillin 1975; Deschamps and Phulpin 1980). Recent results of McClain (1981; McClain *et al.* 1983) demonstrate that the use of two or three channels allows values of global s.s.t. to be inferred which are well within 1 K of *in situ* measurements.

The results presented here were obtained from the AVHRR/2 instrument. This has an instantaneous field of view (at the sub-satellite point) of 1·1 km which scans across the swath perpendicular to the ground track of the satellite. The data are oversampled with the result that at the sub-satellite point the distance between pixel centres is about 0·8 km. The satellite is in a sun-synchronous, near-polar orbit, passing over at local times of approximately 1430 h and 0230 h. The great majority of the comparisons between satellite and ship measurements are from the north-eastern Atlantic and use full resolution satellite data (Local Area Coverage data, LAC), but a few tropical Atlantic comparisons, where the atmospheric absorption can be much higher, are also included using sub-sampled 4 km resolution AVHRR/2 data (Global Area Coverage data, GAC).

2. Atmospheric simulations

The terrestrial radiation received by a space-borne radiometer directed at the earth consists principally (over the ocean) of thermal emission from the sea surface and atmosphere modified by absorption and scattering by the atmosphere. In addition solar radiation can be reflected or scattered into the field of view by the sea surface and clouds. In order to interpret the measurements in space of terrestrial radiance it is necessary to analyse the influence of the atmosphere on the upwelling radiation. This section describes numerical simulations which result in expressions for deriving s.s.t. from the radiance measurements.

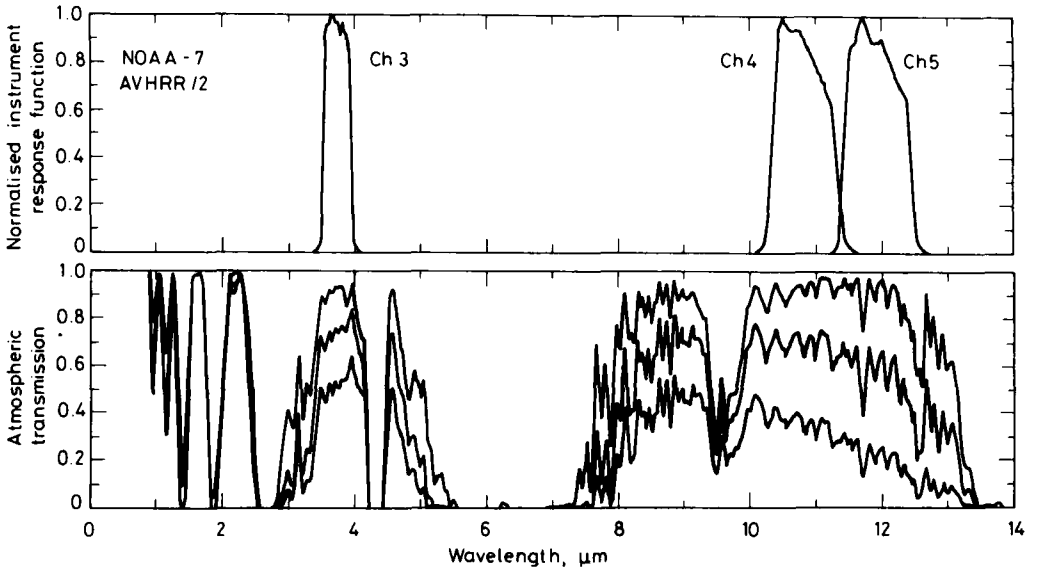

Figure 1. Theoretical spectra of atmospheric transmission, at nadir, in the infrared region at 1 to 14 μm wavelength. The three spectra correspond to different amounts of precipitable water (7 mm — polar; 29 mm — temperate; 54 mm — tropical). Aerosol effects are not included in these spectra. The response functions of channel 3 (3·7 μm), channel 4 (11 μm) and channel 5 (12 μm) of the AVHRR/2 on the NOAA-7 satellite are shown, and indicate the 'atmospheric windows' used. The different dependence of atmospheric transmission on water vapour amounts in each 'window' permits an estimate of the atmospheric effect in s.s.t. measurement by multichannel methods.

TABLE 1. SOURCES OF DATA TO COMPUTE MOLECULAR ABSORPTION

Atmospheric constituents considered	H_2O, CO_2, O_3, N_2O, CH_4
Centre frequencies Line intensities Air-broadened line widths	AFGL compilation, Rothman (1981)
Line shape	Gross (1955)
Criterion for selecting lines	At each frequency, lines within 20 cm^{-1} were included
Total number of lines considered	7170 (3·7 μm channel) 908 (11 μm channel) 925 (12 μm channel)
Frequency resolution	0·04 cm^{-1}
Model for water 3·5–4·2 μm: vapour continuum 10–13 μm:	Barton (1981) } ρ^2 dependence; Bohlander (1979) } exponential T dependence

Theoretical spectra of atmospheric transmission in the wavelength interval 1 to 14 μm are shown in Fig. 1. These were produced using line-by-line calculations of absorption with the model set out in Table 1. The three spectra correspond to different amounts of atmospheric water vapour (7, 29, 54 mm precipitable water) and representative profiles of temperature and they illustrate the behaviour of the 'atmospheric windows' at 3·7, 11 and 12 μm, where the satellite measurements are made. The response functions of the infrared (channels 3, 4 and 5) of the AVHRR/2 instrument are also shown. The absorption in the 3·7 μm window is dominated by CO_2 and, although it has a lower transmission than the others in the case of the dry (polar) atmosphere, it is less variable and significantly more transmissive for moist (tropical) atmospheres. In addition, the 3·7 μm channel has much higher radiometric sensitivity than the 11 and 12 μm channels; for example, the radiated power from a black body at 3·7 μm has a temperature dependence at 300 K of $\sim T^{13}$ compared with $\sim T^5$ at 11 μm. However, while the levels of solar radiation scattered and reflected into the field of view are negligible when compared with the power emitted at terrestrial temperatures in the 11 and 12 μm channels, the terrestrial emission at 3·7 μm is so low that the scattered solar radiation can make a significant contribution to the daytime signal in this channel, often rendering it unusable for s.s.t. determination. Furthermore, simulations show that measurements in the 3·7 μm window can be appreciably degraded by the presence of atmospheric aerosols.

The model used to derive the theoretical spectra in Fig. 1 is the basis of the numerical simulations used to obtain relationships between s.s.t. and the measurements at two or three channels. Calculations have been carried out which predict the brightness temperature of the surface and atmosphere seen from space in each channel for a range of different atmospheric conditions and s.s.t.s.

The particular atmospheric conditions used for these calculations were taken from a representative set of 61 radiosonde ascents recorded over the North Atlantic and another of 39 ascents over tropical oceans. These are subsets of a larger number of soundings assembled by NOAA. Within geophysical limitations, one is at liberty to chose the value of the s.s.t. under each of the atmospheric profiles. In the calculations for this study each profile was assigned five s.s.t. values, these being the surface air temperature of each sounding, T_0, and $(T_0 \pm 2)$K and $(T_0 \pm 4)$K. Line-by-line calculations of radiance over the frequency range of the response functions of the three channels were made over the range of vertical water vapour and temperature distributions recorded by the soundings. The radiation emitted from the sea, and emitted by the atmosphere — the latter including radiation emitted in the direction of the satellite and that reflected

from the sea surface — all contributed to the computed brightness temperatures. The reflectivity of the sea surface, assumed to be smooth, was computed by using the Fresnel equations with the refractive index of water at the frequency of interest. Simulations showed that using the 'smooth sea' reflectivity values gave rise to errors in brightness temperatures of less than 0·1 K, in all AVHRR/2 infrared channels, for surface wind speeds of less than 25 m s^{-1} and for airmasses of less than 1·5. For the largest airmass value used in this study, 2·0, the maximum error was predicted to be 0·2 K up to a surface wind speed of 15 m s^{-1}.

In order to estimate s.s.t. using a radiometer with N channels, measurements are made of the terrestrial radiance R_i, $i = 1, \ldots, N$, in the N wavebands centred at λ_i. The radiances are related to temperature by the Planck function and, for given atmospheric conditions, the surface temperature, $\tilde{T}_s$, can be expressed as $\tilde{T}_s = f(T_1, T_2, \ldots, T_N)$ where f is a complicated function, dependent on atmospheric structure, surface properties, satellite zenith angle, etc. For practical purposes the expression is approximated by

$$T_s = a_0 + \sum_{i=1}^{N} a_i T_i$$

where T_s is the retrieved surface temperature.

Prabhakara *et al.* (1974) proposed that measurements at two (or three) distinct wavebands in the 10–13 μm atmospheric window could be used to determine the atmospheric effect and hence s.s.t. Since, during the day, the 3·7 μm channel brightness temperature has an appreciable contribution from scattered and reflected solar radiation (Takashima and Takayama 1981), only the 11 and 12 μm channels will be used for daytime data, leading to the 'split window' regression relationship, which can take the form

$$T_s = C_0(\alpha) + C_1(\alpha)T_{11} + C_2(\alpha)T_{12} \tag{1}$$

where T_{11} is the 11 μm brightness temperature (K), T_{12} the 12 μm brightness temperature (K), and α is the airmass (Kondratyev 1969).

As Eq. (1) is a linear approximation, the coefficients C_0, C_1, C_2 are functions of the airmass α (which is related to the satellite zenith angle, ψ, by $\alpha \approx \sec \psi$ for $\psi < 60°$). In order to approximate the complicated airmass dependence of the coefficients, simulations were carried out for five airmasses: 1·0, 1·25, 1·5, 1·75 and 2·0, corresponding to satellite zenith angles of 0°, 36·87°, 48·19°, 55·15° and 60·0°. Sea surface temperatures corresponding to intermediate airmasses are obtained by linear interpolation. The coefficients calculated for the North Atlantic are given in Table 2 and for the tropics in Table 3.

The airmass dependence has not been specifically taken into account in previous published derivations of the multichannel algorithms, and the consequence of this omission is shown in Fig. 2(b). The simulations show that errors of up to 2 K in derived s.s.t. can arise if airmass dependence is neglected. A real case of a cloud-free area of sea surface (50 × 50 pixels) recorded through an airmass of 1·32 ± 0·05 by AVHRR/2 was also considered. Using coefficients for unit airmass resulted in a mean s.s.t. of 17·0 °C whereas a value of 17·2 °C was derived using the scheme in which airmass dependence was taken into account. Figure 2(b) shows that this difference is in good agreement with the simulations.

At night all three infrared channels of AVHRR/2 can be used for s.s.t. determinations over cloud-free areas. The 'triple window' regression relationship takes the form

$$T_s = C_0'(\alpha) + C_1'(\alpha)T_{11} + C_2'(\alpha)T_{12} + C_3'(\alpha)T_{3\cdot7}. \tag{2}$$

TABLE 2. SPLIT WINDOW REGRESSION COEFFICIENTS FOR 61 N. ATLANTIC PROFILES

Airmass α	C_0 (K)	C_1	C_2	Expected r.m.s. uncertainty (K)
1·0	−0·334	2·6710	−1·6689	0·07
1·25	0·246	2·8478	−1·8479	0·08
1·50	−0·017	2·9610	−1·9597	0·09
1·75	−1·503	3·0011	−1·9932	0·11
2·0	−5·595	2·9038	−1·8795	0·14

TRIPLE WINDOW REGRESSION COEFFICIENTS FOR 61 N. ATLANTIC PROFILES

Airmass α	C'_0 (K)	C'_1	C'_2	C'_3	Expected r.m.s. uncertainty (K)
1·0	−1·022	2·0732	−1·5247	0·4572	0·06
1·25	−0·585	2·1948	−1·6830	0·4924	0·07
1·50	−0·793	2·1891	−1·7862	0·6027	0·08
1·75	−2·337	2·1629	−1·8252	0·6747	0·09
2·0	−6·912	2·1129	−1·7144	0·6319	0·13

Note. All coefficients assume a noise-equivalent temperature difference of 0·02 K for AVHRR/2.

TABLE 3. SPLIT WINDOW COEFFICIENTS FOR 39 TROPICAL PROFILES

Airmass α	C_0 (K)	C_1	C_2	Expected r.m.s. uncertainty (K)
1·0	−17.1817	3·9078	−2·8524	0·36
1·25	−24.7688	4·2469	−3·1667	0·48
1·50	−33·6119	4·5808	−3·4711	0·61
1·75	−44·2232	4·8849	−3·7391	0·74
2·0	−54.3673	5·1959	−4·0156	0·88

TRIPLE WINDOW REGRESSION COEFFICIENTS FOR 39 TROPICAL PROFILES

Airmass α	C'_0 (K)	C'_1	C'_2	C'_3	Expected r.m.s. uncertainty (K)
1·0	−9·523	−0·1244	−0·7228	1·8854	0·13
1·25	−13·206	−0·4912	−0·5736	2·1173	0·17
1·50	−17·326	−0·8334	−0·4337	2·3356	0·22
1·75	−21·579	−1·1938	−0·2815	2·5607	0·28
2·0	−26·785	−1·4673	−0·1733	2·7463	0·35

Note. All coefficients assume a noise-equivalent temperature difference of 0·02 K for AVHRR/2.

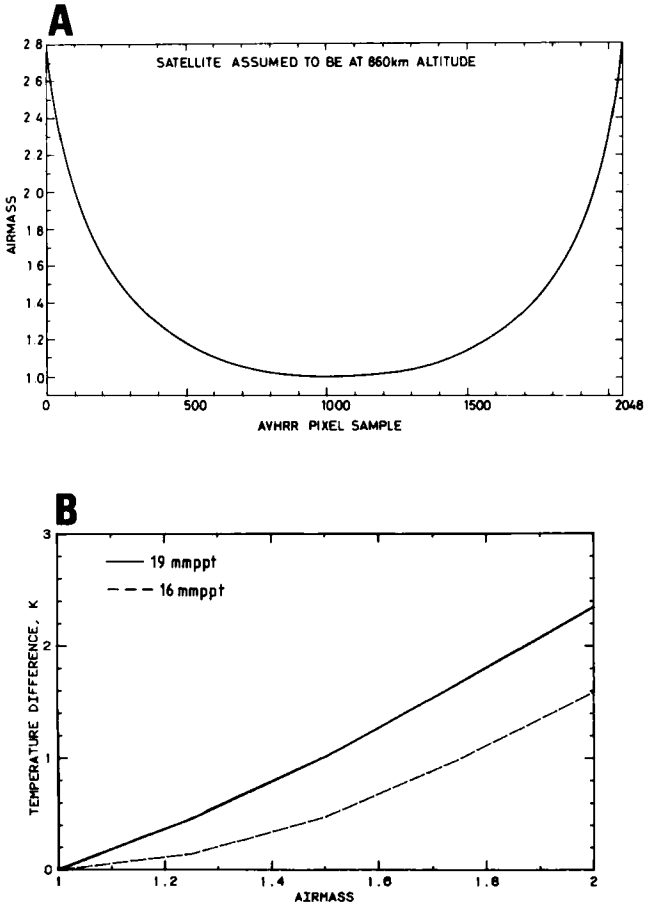

Figure 2. (A) The change in airmass across an AVHRR/2 scan. (B) The effect of neglecting airmass dependence. The difference between s.s.t. derived from the coefficients in Table 2 and from only the coefficients for unit airmass. Two North Atlantic profiles were used with 16 mm (dashed) and 19 mm (solid) precipitable water. The split window (11, 12 μm) coefficients were used.

These coefficients are listed in the second part of Tables 2 and 3.

The coefficients were derived by multiple linear regression of the black body temperatures in the 2 or 3 channels, computed using the 61 North Atlantic and 39 tropical profiles (including the effects of instrumental noise), against the s.s.t. values associated with each profile. This minimizes the variance between the 'true' s.s.t. and the s.s.t. calculated using Eqs. (1) or (2). Variability in the temperature and water vapour structure of the profiles and the simulated effect of instrumental noise give rise to a residual uncertainty in the s.s.t. computed using the coefficients from Tables 2 and 3 in Eqs. (1) and (2). The residual uncertainty, and the values of coefficients themselves, are strongly dependent on the inherent noise level of the signal caused by instrumental effects, which must be taken into account when deriving the regression relationship.

The noise-equivalent temperature difference (NEΔT) of the AVHRR/2 infrared channels is specified as 0·12 K (Lauritson *et al.* 1979). Using this value in the simulations gives the expected r.m.s. uncertainty in a single pixel split window and triple window s.s.t. retrieval of 0·3 K and 0·2 K for the North Atlantic simulations, and 0·8 K and 0·3 K

for the tropics. When s.s.t. is measured over a larger area, however, the uncertainty in the average value is much smaller. It was found that no significant improvement in the residual uncertainty was to be gained in the simulations by reducing, by spatial averaging, the effective NEΔT below 0·02 K. Using this value with 61 North Atlantic simulations gave expected r.m.s. uncertainties of 0·08 K for the split window and 0·07 K for the triple window. The 39 tropical split window simulations gave an expected r.m.s. uncertainty of 0·48 K, which is considerably more than for the North Atlantic because of much larger water amounts and hence higher atmospheric absorption at wavelengths of 11 and 12 μm. The triple window simulations for the tropics gave an r.m.s. uncertainty of 0·17 K, which is a significant improvement over the split window uncertainty, indicating that, theoretically at least, the 3·7 μm channel is essential for the retrieval of tropical s.s.t. This is due to the lower contribution to atmospheric absorption by water vapour at 3·7 μm. The necessity of 3·7 μm measurements at temperate latitudes is not demonstrated by these simulations. All r.m.s. uncertainties given above are for the case where the sea surface is viewed through an airmass of 1·25. The complete list of expected r.m.s. uncertainties, for clear conditions and airmasses up to 2·0, are given in Tables 2 and 3. These values give an idea of the best accuracy theoretically obtainable from the two or three wavelength channels of AVHRR/2.

3. Measurements of s.s.t. from NOAA-7

(*a*) *Pixel location for NOAA-7*

The geographical positions of the pixels in an image can be calculated using orbit elements. These were supplied by the University of Aston Earth Satellite Research Unit. The track of the sub-satellite point was calculated as a function of time following the algorithms of Ruff and Gruber (1975). The AVHRR/2 data tapes record the time of each scan to millisecond accuracy and, using the description of the instrument provided by NOAA (Schwalb 1978), the geographical coordinates of each pixel can be calculated using simple geometry.

By comparing the derived latitudes and longitudes of headlands in the image with their known positions, the accuracy of the method can be checked. Errors of up to a few tens of kilometres have been found, but generally they are less than this. The principal sources of error appear to be uncertainties in both the time and the longitude of the ascending node (i.e. when and where the satellite crossed the equator on the northbound part of its orbit). In order to remove cloud contamination (see below) many pixels (2500) are used to give a single s.s.t., so that when comparing this with a given *in situ* measurement at a given geographical point, even the largest errors in pixel location should be accommodated in the area encompassed by the satellite data.

(*b*) *Calibration of AVHRR/2*

The AVHRR/2 data are received from the satellite in the form of radiometer counts between 0 and 1023, which are related to the incident radiance within each channel. Part of each scan includes a space view and a view of the in-flight calibration target, a black body, whose temperature is monitored by four platinum resistance thermometers. The calibration procedure described by Lauritson *et al.* (1979) was followed. The mean temperature of the four thermometers on the calibration target was used to compute the emitted radiances of the target at each of 60 wavenumbers, uniformly spaced across the width of each infrared channel, at which the normalized response function of each channel is defined. The resultant calibration target radiance was used, together with the

space-view measurements, to give a linear relationship between counts and radiance. This was adjusted using a first-order approximation to the nonlinear response of the detectors of the 11 and 12 μm channels (the 'non-zero space radiance correction') and used to convert the counts to radiance values. The infrared radiances then were converted to brightness temperature, again using the normalized response function across the full width of the channel. The residual temperature errors, caused by the nonlinearity of the channels 4 and 5 detectors, are tabulated by Lauritson *et al.* These errors are zero at $T = 285$ K, but range from $-0{\cdot}4$ K and $-0{\cdot}2$ K for channels 4 and 5 at $T = 275$ K to $+0{\cdot}5$ K and $+0{\cdot}3$ K at $T = 295$ K. The tabulated values were linearly interpolated to provide the final corrections to the measured brightness temperatures at the 11 and 12 μm channels.

The brightness temperatures resulting from this calibration procedure, T_{11}, T_{12} and $T_{3{\cdot}7}$, are used in Eqs. (1) and (2) to give sea surface temperature.

(*c*) *Removal of cloud contamination*

When measured in space, infrared radiance from the sea surface is often contaminated by that from clouds. Since clouds can be very much colder than the sea, even undetected cloud cover amounting to just a few percent of the area of a given pixel can introduce large errors in the retrieved s.s.t. It is therefore necessary to identify all pixels contaminated by clouds and eliminate them from the s.s.t. retrieval process. There are several techniques for cloud identification, each with varying success in different situations, and the choice of technique can be as important to accurate s.s.t. determination as the choice of multichannel algorithm.

Three methods for the detection of cloud-contaminated pixels were investigated for this study. The first two, the truncated normal distribution technique and the spatial coherence technique, rely on the statistical properties of the infrared radiance in a group of pixels; the third makes use of the high albedo of clouds at shorter wavelengths. The size of the array of pixels for the statistical treatment, 50×50, is based on the results of an earlier study (Harris *et al.* 1981).

The Truncated Normal Distribution (TND) technique (Crosby and Glasser 1978; Harris *et al.* 1981) can be used to determine the s.s.t. in cloudy images by night and day, including areas of sunglint. The principle is to fit a Gaussian curve to the warm edge of a histogram of the infrared radiances, or brightness temperatures, and infer a cloud-free value from the peak value of the fitted curve. In regions of s.s.t. variability, the s.s.t. value derived with this technique will be biased towards the warmest values in the area.

The other technique using emitted radiances is based on the Spatial Coherence method described by Coakley and Bretherton (1982). Over cloud-free sea surface the local standard deviation in brightness temperature should be small, whereas for cloud-contaminated pixels the local standard deviation will be much larger. However, in some cases cloud-contaminated pixels may not be identified since uniform low cloud tops can also have low local standard deviations. For this study sub-arrays of 3×3 pixels of AVHRR/2 11 μm (channel 4) data were used, and only those which had a local standard deviation of less than 0·1 K were included in the mean s.s.t. of the 50×50 array. This local standard deviation threshold was determined by varying its value for a few test cases and observing when pixels within the cloud-free peak of the histogram began to be rejected. This spatial coherence method results in most of the cloud-contaminated pixels being removed.

During the day, scattered and reflected solar radiation makes a strong contribution to the total upwelling radiance in channels 1 (visible) and 2 (near infrared) of AVHRR/2. Over the cloud-free ocean, outside sunglint regions, the reflected radiance

TABLE 4. DATA FROM RESEARCH AND WEATHER SHIPS

Ship	Area	Measurement	Position (Fig. 3)
Cumulus	N. Sea, Norwegian Sea	WMO	C
Discovery	N. Atlantic, N. Sea, Gulf of Guinea	C	D
Frederick Russell	English Channel, N. Sea	P	F
Meteor	N. Atlantic	C	M
Noord Hinder	N. Sea	WMO	N
Poseidon	N. Atlantic	WMO	P
Tyro	N. Sea	WMO, P	Y
Tydeman	N. Atlantic	C	T

Measurement key: WMO — bucket thermometers to WMO schedule; P — shallowest value in a measurement of the temperature profile; C — continuous near-surface temperature.

is low, whereas clouds give rise to high reflected radiances. This allows a simple threshold technique to be employed so that only radiances from cloud-free pixels are used for s.s.t. measurements.

(*d*) *Ship–satellite comparisons*

In order to check the validity of the simulations, the s.s.t. values derived from satellite data using Eq. (1) are compared with *in situ* measurements from ships. The data used here are mostly from research ships. The vessels and measurement techniques are listed in Table 4. Ship measurements have been restricted to those made within two and a half hours of the satellite overpass, to reduce uncertainties due to the diurnal thermocline (Ostapoff and Worthem 1974; Saunders *et al.* 1982). Under conditions of high insolation and low surface wind speeds this can cause changes of s.s.t. of a few tenths of a degree in a few hours. Measurements from ships using 'continuous' monitoring equipment have been favoured as the time interval between ship and satellite measurements is then reduced to a few minutes. The comparisons were in the North Atlantic between latitudes 37°N and 65°N, and in the tropical Atlantic at a latitude of 5°S. The positions of the vessels at the time of each comparison are shown in Fig. 3.

During the period which was chosen to compare satellite s.s.t. measurements with those from ships (July to September 1981) the occasions when *in situ* measurements coincided with the reception of data from the satellite were generally restricted to the daytime. The satellite s.s.t. values are from cloud-free and from partly cloud-obscured arrays but are confined to cases with airmasses of less than 2·0, corresponding to a swath width of 2280 km. This gives complete coverage (i.e. overlapping swaths from successive orbits) polewards of latitude 32°. To obtain global coverage from NOAA-7 a swath width of 2840 km would be necessary, corresponding to airmass values of about 2·5. However, as the airmass increases, so does the angle of emission of the radiation that is detected in space. Calculations have shown that at emission angles greater than 60°, the emissivity begins to show strong dependence on the angle of emission, which, because of the effect of waves tilting the sea surface, is also a function of wind speed (Cox and Munk 1954). Because of this added complexity, and possible source of error, the results of the simulation model were not applied to situations with airmasses greater than 2·0. This restriction means a loss of about 100 pixels at each edge of the scan.

Apart from exceptional cases of specular reflection from a calm sea, the sunglint contamination which occurs in the 3·7 μm radiances is absent from data measured in the

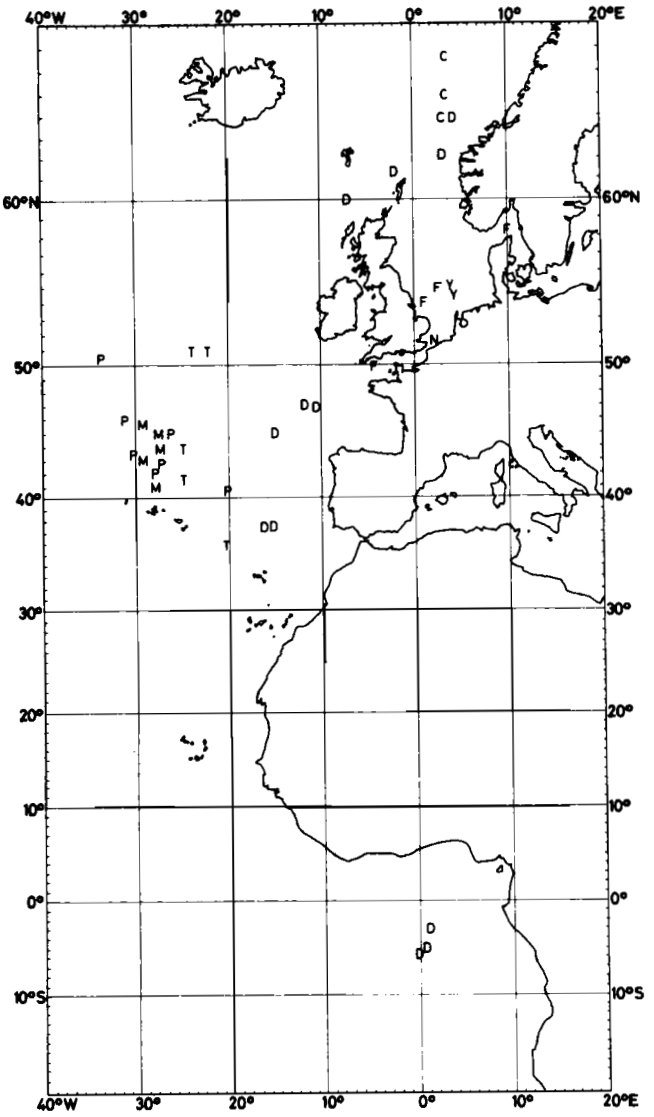

Figure 3. The position of the research vessels at the time of each ship–satellite comparison. For explanation of code letters and method of observation, see Table 4. Although the data are seasonally localized, being from July to September 1981, the large latitudinal range encompassed is apparent.

split window (11 and 12 μm). The comparison between ship and satellite-derived s.s.t. values is shown in Fig. 4. The North Atlantic points are mean values of cloud-free pixels from a 50 × 50 array (41 × 55 km at nadir, 130 × 56 km at airmass 2·0) nominally centred at the position of the *in situ* measurements. The tropical points are taken from 14 × 14 GAC pixels (Global Area Coverage data; Schwalb 1978) corresponding to areas of 56 × 44 km at nadir and 176 × 49 km at airmass 2·0. Pixels contaminated with cloud radiances were identified by applying the spatial coherence technique to the 11 μm brightness temperatures followed by a visible channel threshold discriminator, in which the threshold level is found for each case from the histogram of the channel 1 counts

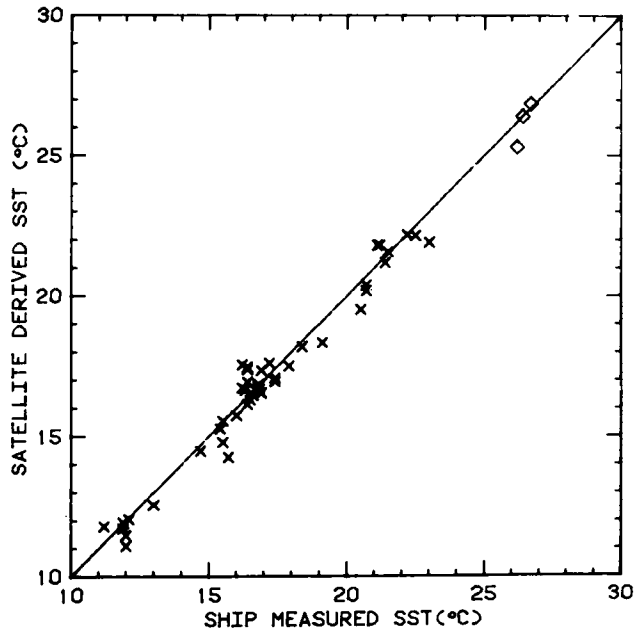

Figure 4. A comparison of 52 s.s.t. values measured by ships and the corresponding values derived from AVHRR/2 using the split window relationship over a 50×50 pixel array in which the ship is centred. All ship observations were within two and a half hours of the satellite overpass. (X: North Atlantic points; ◇: tropical points.)

(hereafter referred to as the SCV method). The 49 North Atlantic points in Fig. 4 have differences between satellite and ship s.s.t. (satellite − ship) with a mean and standard deviation of −0·13 K and 0·58 K. Only 10% of these points show a discrepancy of more than 1·0 K and 63% are closer than 0·5 K.

To investigate the reasons for the discrepancies between ship and satellite s.s.t. the differences are plotted as a function of airmass in Fig. 5(a) and as a function of the number of cloud-free pixels in Fig. 5(b). These figures include only North Atlantic points.

The airmass plot shows that no significant correlation was present. This indicates that the airmass dependence of the atmospheric correction is well modelled in the numerical simulations.

The differences appear to be independent of the cloud amount (number of cloud-free pixels in array) as shown in Fig. 5(b). This increases our confidence in the SCV cloud removal technique as otherwise the differences would increase with cloudiness owing to the effect of contaminated pixels.

(*e*) *The use of a 'global' retrieval scheme*

The use of regionally distinct retrieval coefficients (e.g. tropical, North Atlantic, etc.) relies on being able correctly to choose the set of coefficients appropriate to the atmosphere through which the satellite measurement was made. This requires identifying the possibly indistinct boundaries of atmospheric regimes, and using some mechanism of smoothly changing from one set of coefficients to another without introducing discontinuities in the retrieved s.s.t. near these boundaries. In an operational scheme, intended to retrieve s.s.t. on a global basis, this could be a serious problem but which is avoided by using a single set of globally applicable coefficients. This inevitably leads

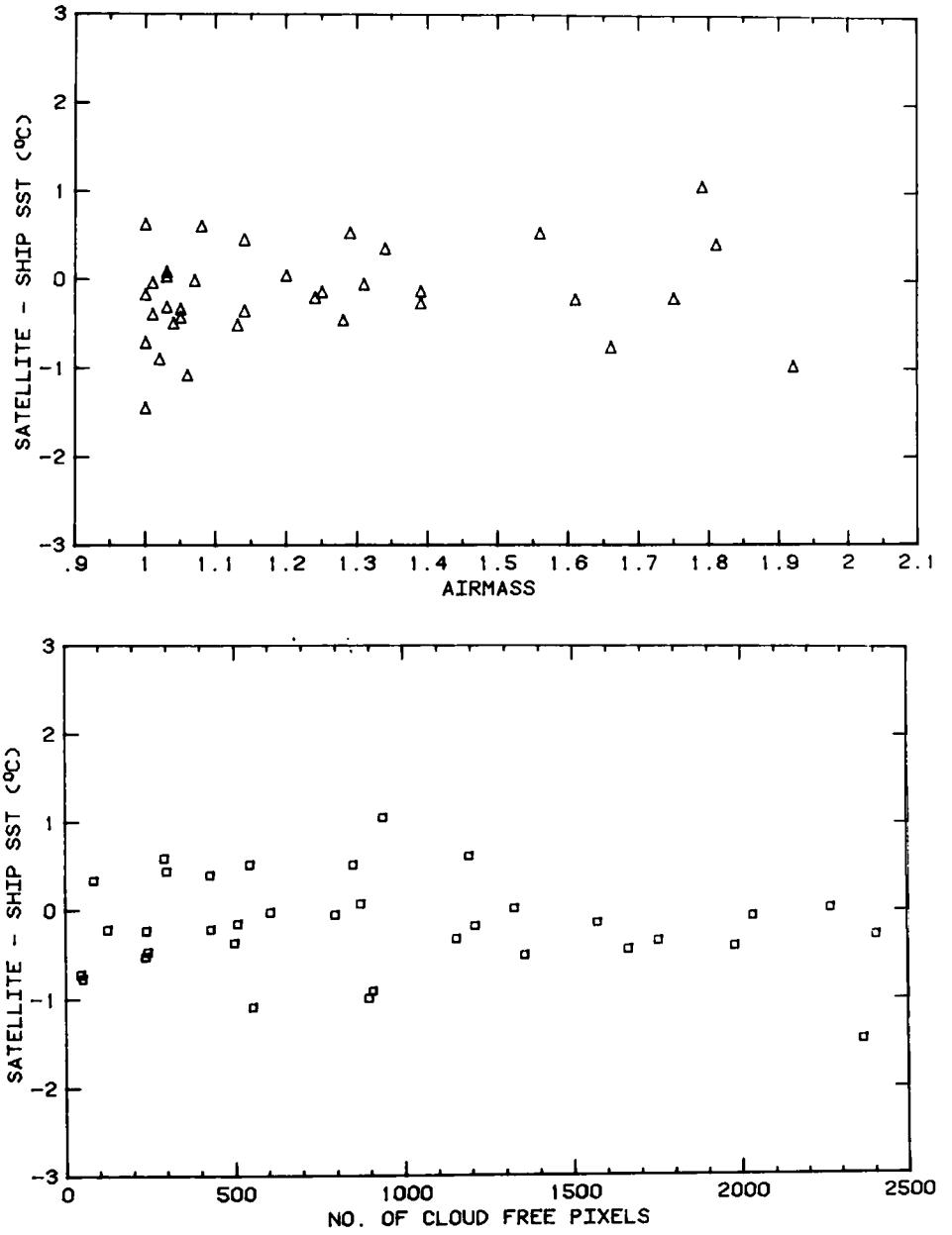

Figure 5. Differences between ship and satellite s.s.t. values plotted as a function of airmass and as a function of the number of cloud-free pixels in the 50×50 array.

to a degradation in the quality of the s.s.t. retrievals in any particular area where the global diversity of atmospheric variability would not occur. The severity of the degradation was investigated by recalculating the North Atlantic s.s.t. values, described above, using 'global' coefficients derived from the combined set of North Atlantic and tropical atmospheric profiles. The differences between satellite and ship values have a mean and standard deviation of 0·25 K and 0·62 K compared with −0·13 K and 0·58 K when North

Atlantic coefficients are used. Thus, in this case, the benefit of using regional coefficients is to reduce significantly the mean difference between satellite and ship measurements, but to leave the scatter practically unchanged.

(*f*) *Comparison with NESDIS m.c.s.s.t. algorithms*

Since late 1981 the National Environment Satellite, Data and Information Service (NESDIS) of the National Oceanic Atmospheric Administration (NOAA) in the United States has been issuing measurements of s.s.t. from AVHRR/2 on an operational basis. The procedure involved in producing the multi-channel sea surface temperature (m.c.s.s.t.) is described by McClain *et al.* (1983).

For daytime measurements a cloud elimination scheme similar to the SCV method, described above, is used, with an additional test which requires the differences between measured brightness temperatures in the infrared channels to be below given values. This is intended to identify cloud contamination which has slipped through the visible albedo and spatial uniformity tests (i.e. thin, or uniform, or certain types of sub-pixel-size, cloud fields).

The cloud-free pixels are then processed using the relationship

$$\text{m.c.s.s.t.} = -10{\cdot}77 + 1{\cdot}035T_{11} + 3{\cdot}046(T_{11} - T_{12}). \tag{3}$$

The coefficients were derived by McClain *et al.* using a set of 59 cloud-free radiosonde observations in conjunction with an atmospheric transmission model (McClain 1981), with a subsequent empirical bias correction. The dependence on the airmass through which the measurements are made is not explicitly treated, but this is justified by restricting the m.c.s.s.t. retrievals to a subswath defined by satellite zenith angles less than 45°. A scheme to extend the range of retrievals beyond this is currently being studied (McClain, 1983, private communication), using an expression of the form:

$$\text{m.c.s.s.t}'. = a + T_{11} + b(T_{11} - T_{12}) + c(T_{11} - T_{12})(\sec\psi - 1) \tag{4}$$

where the last term is an approximation of the effect of increasing airmass with increasing satellite zenith angle, ψ.

The set of near-coincident ship and satellite measurements described in the previous section was also used to compare the m.c.s.s.t. values derived from Eqs. (3) and (4) with *in situ* temperatures. The coefficients in Eq. (4) are given by McClain and Walton (personal communication, 1983) as $a = -0{\cdot}14$ K, $b = 2{\cdot}346$, $c = 0{\cdot}655$. To allow a direct comparison with the results obtained using the algorithms developed in this study, the m.c.s.s.t. retrievals were preceded with the same cloud elimination scheme (SCV) described earlier, and not that given by McClain *et al.* The results of 39 comparisons with the ship measurements are shown in Table 5.

4. Split window — triple window comparisons

As stated in section 2, low values of atmospheric attenuation and the nature of the Planck function make the 3·7 μm channel very advantageous to use in tropical regions. Because of the absence of an adequate quantity of research ship data, it has not been possible to verify empirically the triple window algorithm in the tropics. Furthermore, the problem of sunglint contamination of available daytime data and the absence of nighttime data also prevented a verification at higher latitudes. It is possible, however, to check for self consistency between the split window s.s.t. values derived from Eq. (1) and the triple window s.s.t. values from Eq. (2).

TABLE 5. STATISTICS OF S.S.T. RETRIEVALS (SATELLITE − SHIP)

Method	This study: (N. Atlantic coefficients) Eq. (1)		This study: ('global' coefficients) Eq. (1)		m.c.s.s.t. (NOAA-NESDIS): Eq. (3)		m.c.s.s.t.' (with secant approximation): Eq. (4)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
SCV	−0·104	0·533	0·256	0·617	0·285	0·682	0·181	0·538
TND	0·237	0·589	0·750	0·670	0·807	0·713	0·688	0·614

AVHRR/2 split window (11 and 12 μm) s.s.t. statistics of 39 cases where TND fit successful. SCV ≡ cloud removal by spatial coherence and visible threshold techniques. TND ≡ cloud removal by truncated normal distribution technique.

Two nighttime passes of NOAA-7 on 17 and 25 September 1981 were selected, with low temperatures measured over the Norwegian Sea and higher temperatures over the Mediterranean Sea. Cloud-contaminated pixels were identified using the spatial coherence method alone, and average temperatures from cloud-free pixels in 50 × 50 arrays were used to give s.s.t. from the triple window and split window algorithms. The results are plotted in Fig. 6. The small amount of scatter suggests that the 3·7 μm channel is not seriously affected by aerosols for these atmospheric conditions. The triple window appears to give s.s.t. values slightly higher than the split window (mean and standard deviation of difference: 0·23 K, 0·15 K) but with the data available it is not possible to

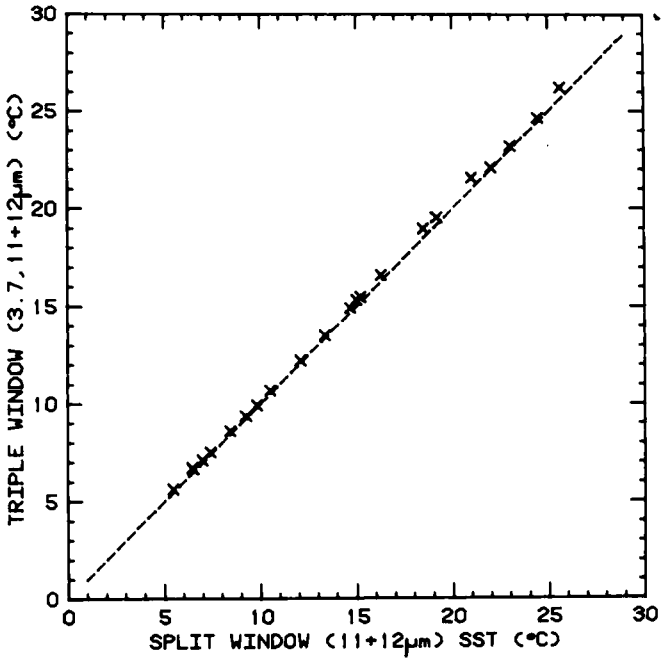

Figure 6. A comparison of s.s.t.s at night for 17 and 25 September 1981, derived using the split window and the triple window relationships.

determine which algorithm gives less error. The difference between the two shows no significant correlation with airmass, which makes it unlikely that the modelling is in error. A small error in the calibration may be responsible, caused, say, by uncertainties in the emissivity of the black body at 3·7 μm.

5. Discussion

(*a*) *Possible sources of error*

The factors which contribute to the discrepancies between ship and satellite measurements of s.s.t. can be grouped into three classes: atmospheric; instrumental (including the accuracy of *in situ* measurements); procedural (i.e. errors arising from the method of comparison). The last group includes uncertainty in co-registration of ship and satellite measurements of s.s.t., and near-surface vertical temperature gradients, caused by the skin effect and the diurnal thermocline, which result in temperature differences between the sea surface and at the depth of the *in situ* measurement. Ultimately, uncertainties in the atmospheric factors (temperature deficit and cloud contamination) will limit the accuracy to which s.s.t., averaged over large areas, can be measured from space. However, it is clear that procedural factors contribute significantly to the discrepancies found in this study.

The absolute accuracy of the satellite measurement of s.s.t. is critically dependent on the in-flight calibration. There is evidence that the on-board black-body calibration target is not ideal, as the temperatures measured simultaneously by the four thermometers can differ by more than 1·5 K (assuming the thermometer calibration to be correct) and the mean of these temperatures changes with time. It is presumed that the target mean temperature and temperature gradients are dependent on the thermal condition of the instrument and consequently on the time elapsed since leaving eclipse (or entering eclipse for nighttime data). The temperature variability of the target, both spatial and temporal, could result in the radiation temperature of the target, or that part of the target used in the calibration procedure, being different from the value given by the mean temperature of the four thermometers. Since the error in s.s.t. is equal to the calibration error in each channel multiplied by the retrieval coefficients, a calibration error affecting all channels equally leads to a comparable error in s.s.t. If, however, a calibration error affects the channels differently, amplification of the error in s.s.t. can occur. Such a case might arise if the infrared channels were not exactly co-registered and temperature gradients existed on the target surface, or if there were a post-launch change in the detector characteristics, or a post-launch change in the emissivity of the black-body target. For example, a 0·1% change in the emissivity of the target at channel 4 relative to channel 5 would cause an error of 0·24 K in s.s.t., whereas the same emissivity change affecting both channels equally would result in an s.s.t. error of 0·08 K. The contribution of these calibration error sources to the discrepancies between the satellite and *in situ* measurements cannot be confidently quantified.

Anomalous atmospheric conditions (in the sense of being very different from those defined by the set of atmospheric profiles used in the simulations) at the time of the satellite measurements, or defects in the numerical simulations, would produce discrepancies with the *in situ* temperatures that increase with increasing airmass. The observed lack of airmass dependence in the admittedly small number of cases treated here (Fig. 5(a)) indicates that these two factors do not contribute significantly to the scatter in Fig. 4. However, the mean difference between satellite and *in situ* temperatures could be related to the fact that the North Atlantic profiles used in the simulations are drier

than would be expected for the time of year in which the comparisons were made (P. K. Taylor, 1983; private communication).

Failure to identify and remove cloud-contaminated pixels would lead to a lowering of the satellite s.s.t., and while this may contribute to the bias in Fig. 4, it, by itself, cannot explain the scatter, which includes many cases where the satellite s.s.t. is higher than the *in situ* measurement.

These cases could be caused by the diurnal thermocline warming the sea surface relative to the *in situ* measurements, which are taken at a depth ranging from a few centimetres to about one metre. By the time of the satellite overpass (~1430 h local time), temperature differences between the surface and a depth of one metre of several tenths of a Kelvin could have developed on calm, sunny days. On the other hand, cooling of the surface 'skin' of the ocean by heat loss to the atmosphere could lead to *in situ* measurements being say 0·2 to 0·5 K higher than the radiometric 'skin' temperature (Paulson and Simpson 1981). While the diurnal thermocline and the skin effect produce opposite effects, there is no reason for their contributions to cancel, and between them they could account for much of the scatter in Fig. 4.

Spatial variability of s.s.t. within the 50 × 50 pixel array could also account for some of the observed discrepancies. In such cases the point *in situ* measurement may not be representative of the mean cloud-free value. For each North Atlantic point in Fig. 4, the s.s.t. was also calculated using the TND method, using all four multichannel algorithms described above (Table 5). There were six cases where the TND method failed to return an s.s.t. value because of large cloud amounts. The TND s.s.t. was always higher than the corresponding SCV values, and occasionally large differences (>0·5 K) were seen. It is likely that horizontal gradients in s.s.t. caused these as the TND technique will be biased towards the warmer population in a broad distribution. This can be seen in the significantly increased mean differences in the TND results in Table 5. This is an important geophysical limitation to the use of the TND technique. Any effective method for the retrieval of large-scale average values of s.s.t. must be capable of unambiguously detecting and adapting to the presence of s.s.t. gradients. Now, restricting the ship–satellite comparison to the 32 North Atlantic cases without large differences (<0·5 K) between TND and SCV temperatures, the mean and standard deviation remain essentially unchanged, being −0·104 K and 0·517 K. This indicates the performance of the SCV method of cloud clearing is not significantly degraded in the presence of s.s.t. gradients.

A further inevitable source of scatter in Fig. 4 lies in the fact that the *in situ* measurements were made by a variety of different ships using a range of methods. By restricting this study to data of high quality, this source of error is smaller than would be the case had, say, engine intake temperatures also been used. Even so, the absolute accuracy of the mercury in glass thermometers used in bucket measurements is given as ±0·2 K (Meteorological Office 1981) against which hull-mounted thermometers have mean errors of between 0·02 K and 0·15 K (loc. cit.). The *in situ* measurements from profiling devices should be much more accurate, say to 0·01 K given a reliable calibration procedure. However, the ships themselves can disturb the temperature field by amounts comparable to the observed scatter, even when the ship is anchored or free drifting on station (Stevenson 1964). Thus, a significant part of the scatter in the satellite–ship differences must be attributed to the *in situ* data.

(*b*) *Satellite algorithms*

The lower residual bias and standard deviation (Table 5) indicates that the algorithm developed in this study entirely from line-by-line calculations of atmospheric transmission appears to be more accurate when compared with these *in situ* measurements, than the

empirically adjusted algorithm currently used in the operational m.c.s.s.t. retrievals. The m.c.s.s.t. algorithm, however, is designed for global use and the larger discrepancies may result from the fact that the comparisons used here are localized in space and time. The inclusion of the secant approximation in the m.c.s.s.t. scheme significantly improves both scatter and bias, bringing the scatter to the level achieved with our algorithm in which the airmass dependence is treated explicitly.

6. Conclusions

Numerical modelling of the performance of the AVHRR/2 predicts, for the North Atlantic, single pixel accuracy of about 0·3 K in clear air for daytime measurements in the 11 and 12 μm split window, with an increase in accuracy to 0·2 K under ideal conditions at night using the 3·7 μm measurements as well. For the tropics these accuracies decrease, due to higher atmospheric absorption, to 0·8 K in clear, daytime conditions and 0·3 K at night. The simulations also predict that the clear air figures can be better than 0·1 K for NEΔT values lower than 0·5 K. This low value of NEΔT can be achieved in effect by retrieving a spatial average s.s.t. over a large number of pixels, such as is required for many applications needing high absolute accuracy.

Comparison between *in situ* measurements from research ships and s.s.t. values derived from satellite measurements, using algorithms derived entirely from atmospheric simulations without empirical correction, indicates that an r.m.s. accuracy of 0·53 K with very low bias (0·1 K) can be achieved in practice. This result, using data only from the split window at 11 and 12 μm, is comparable to that (0·54 K) obtained by Bernstein (1982) using an empirically derived algorithm on measurements at 3·7 and 11 μm from AVHRR on the NOAA-6 satellite. The use of 3·7 μm data during the day, however, requires very stringent rejection criteria to avoid contamination by reflected solar radiation (sun glint). Using the split window measurements obviates the need to consider the earth–sun–satellite geometry, and the surface wind field, as measurements at this waveband are not susceptible to sun glint contamination. Consequently all of the AVHRR swath can, in principle, be used. This result demonstrates the great potential of the 11–12 μm split window channels on AVHRR/2 to measure s.s.t. on a global basis.

The atmospheric correction was found to depend strongly on the airmass through which the measurement was made. Failure to account properly for this results in increased scatter and bias errors. Corrections using airmass-dependent coefficients or a secant approximation, as proposed by McClain, seem equally good.

The use of the truncated normal distribution (TND) technique to retrieve s.s.t. from pixel arrays containing cloud radiances produced consistently higher temperatures than the spatial coherence and visible threshold method (SCV). There remains much scope for improved methods of identifying cloudy pixels.

A determination of the ultimate limit imposed by atmospheric effects upon s.s.t. measurement accuracy clearly requires improved knowledge of the contributions from the procedural and instrumental factors mentioned above. It is only after these have been fully evaluated that meaningful error budgets for the whole measurement process can be produced. This requires proper estimation of, for example, the effects of the diurnal thermocline, skin effect and spatial variability of s.s.t., as well as more accurate *in situ* measurements. Alternatively a comparison between satellite measurements and a ground based data set derived from surface radiometric measurements would avoid many of the procedural factors altogether.

Several authors have recently published estimates of the accuracy of s.s.t. measurements using a different combination of channels (e.g. Bernstein 1982), or different

retrieval algorithms (e.g. McClain *et al.* 1983), and the values appear to converge at ~0·5 K. These estimates are reached by comparing satellite and *in situ* measurements, and admittedly the values could be lower if 'ground truth' data were available to permit a better comparison than between the spatially averaged, radiometric, satellite measurement and the sub-surface, point, *in situ* temperature. However, in order to achieve a significantly improved atmospheric correction, such as required to give the accuracy necessary for climate research, additional new measurement techniques will have to be used, such as multi-angle scanning, giving a direct indication of the atmospheric correction rather than inferring it from spectral effects.

Acknowledgments

The authors are indebted to E. Paul McClain of NOAA/NESDIS for his valuable comments during the preparation of this paper, and for permission to use unpublished algorithms. The research ship data were supplied by: Nederlands Centrum voor Oceanografische Gegevens at KNMI, de Bilt, Holland: *Cumulus*, *Noord Hinder* and *Tydeman*; Marine Information and Advisory Service, at IOS, Wormley, UK: *Discovery*; R. D. Pingree, Marine Biological Association, Plymouth, UK: *Frederick Russell*; J. Meincke, University of Kiel, FRG: *Meteor*; U. Herrmannsen, University of Kiel, FRG; *Poseidon*; H. M. van Aken, University of Utrecht, Holland: *Tyro*.

The digital tapes of the North Atlantic AVHRR/2 data were purchased from P. E. Baylis of the University of Dundee, and the tropical data from the Satellite Data Service Division of NOAA/NESDIS. The orbital elements used to locate the pixels in geographical coordinates were obtained from C. P. Cooke of the University of Aston.

References

Barton, I. J. 1981 Water vapour absorption in the 3·5–4·2 μm atmospheric window. *Quart. J. R. Met. Soc.*, **107**, 967–972

Bernstein, R. L. 1982 Sea surface temperature estimation using the NOAA-6 satellite Advanced Very High Resolution Radiometer. *J. G. Res.*, **87**, C12, 9455–9465

Bohlander, R. A. 1979 'Spectroscopy of water vapour'. PhD thesis, University of London

Coakley, J. A. and Bretherton, F. P. 1982 Cloud cover from high resolution scanner data: detecting and allowing for partially filled fields of view. *J. G. Res.*, **87**, C7, 4917–4932

Cox, C. S. and Munk, W. H. 1954 Statistics of the sea surface derived from sun glitter. *J. Mar. Res.*, **13**, 198–227

Crosby, D. S. and Glasser, K. S. 1978 Radiance estimates from truncated observations. *J. Appl. Met.*, **17**, 1712–1715

Deschamps, P. Y. and Phulpin, T. 1980 Atmospheric correction of infrared measurements of sea surface temperature using channels at 3·7, 11 and 12 μm. *Bound.-Layer. Met.*, **18**, 131–143

Gross, E. P. 1955 Shape of collision-broadened spectral lines. *Phys. Rev.*, **97**, 395–403

Harris, A. W., Llewellyn-Jones, D. T., Harries, J. E. and Williamson, E. J. 1981 'Use of the histogram technique for infrared measurements of sea surface temperature under conditions of partial cloud cover using data from AVHRR'. Pp. 199–201 in Extended abstracts presented at the symposium on the radiation transfer in the oceans and remote sensing of ocean properties. IAMAP Third Scientific Assembly 17–28 August 1981, Hamburg

Kondratyev, K. Ya. 1969 *Radiation in the atmosphere.* Academic Press, 161–171

Lauritson, L, Nelson, G. J. and Porto, F. W. 1979 'Data extraction and calibration of TIROS-N/NOAA radiometers'. NOAA Tech. Memo NESS 107 44–46

McClain, E. P.	1981	'Multiple atmospheric-window techniques for satellite-derived sea surface temperatures'. Pp. 73–85 in *Oceanography from Space*, J. F. R. Gower (Ed). Plenum Press
McClain, E. P., Pichel, W., Walton, C., Ahmad, Z. and Sutton, J.	1983	Multichannel improvements to satellite-derived global sea surface temperatures. *Advances in Space Research*, **2,** 43–47
McMillin, L. M.	1975	Estimation of sea surface temperature from two infrared window measurements with different absorption. *J. G. Res.*, **80,** 5113–5117
Meteorological Office	1981	*Handbook of Meteorological Instruments* (Second Edition), Vol. 2, HMSO, London
Ostapoff, F. and Worthem, S.	1974	The intradiurnal temperature variation in the upper ocean layer. *Journal of Physical Oceanography*, **4,** 601–612
Paulson, C. A. and Simpson, J. J.	1981	The temperature difference across the cool skin of the ocean. *J. G. Res.*, **86,** C11, 11044–11054
Prabhakara, C., Dalu G. and Kunde, V. G.	1974	Estimation of sea surface temperature from remote sensing in the 11 to 13 μm window region. *ibid.*, **79,** 5039–5044
Rothman, L. S.	1981	AFGL atmospheric absorption line parameters compilation: 1980 version. *Appl. Opt.*, **20,** 791–795
Ruff, I. and Gruber, A.	1975	'Graphical relations between a satellite and a point viewed perpendicular to the satellite velocity vector (side scan)'. NOAA Tech. Memo NESS 65
Saunders, R. W., Ward, N. R., England, C. F. and Hunt, G. E.	1982	Satellite observations of sea surface temperatures around the British Isles. *Bull. Am. Met. Soc.*, **63,** 267–272
Schwalb, A.	1978	'The TIROS-N/NOAA A-G satellite series'. NOAA Tech. Memo NESS 95
Stevenson, R. E.	1964	The influence of a ship on the surrounding air and water temperatures. *J. Appl. Met.*, **3,** 115–118
Takashima, T. and Takayama, Y.	1981	Estimation of sea surface temperature from remote sensing in the 3·7 μm window region. *J. Met. Soc. of Japan*, **59,** 876–890
World Climate Research Programme	1981	'Report of meeting on co-ordination of plans for future satellite observing systems and ocean experiments to be organised within the WCRP, Jan. 1982'. Available from W.M.O. Secretariat