Science and Engineering Research Council
Rutherford Appleton Laboratory
Chilton DIDCOT Oxon OX11 0QX
RAL-93-095

The Role of Satellite Observations of Sea-Surface Temperature in the Detection of Global Change

M R Allen

December 1993

The role of satellite observations of sea-surface temperature in the detection of global change

Myles R. Allen
Rutherford Appleton Laboratory
Chilton, Didcot, OX11 0QX, U.K.
&
Dept. of Physics, University of Oxford,
Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, U.K.
e-mail: atmmra@vax.oxford.ac.uk

November 28, 1993

Acknowledgments

This work was commissioned by the U.K. Department of the Environment as a component of the scientific case for the Advanced Along-Track Scanning Radiometer, and carried out in the Rutherford Appleton Laboratory Space Science Department. I would like to thank Chris Mutlow, John Harries and David Llewellyn-Jones of R.A.L.; Ian Barton of CSIRO, Austrialia; David Anderson of Oxford University; and Phil Jones of the University of East Anglia for helpful comments on this report. Alan Apling, Peter Mallaburn and Vicki Simpson of the D.o.E. were also very helpful. Mike Panter of R.A.L. produced figures 12 to 19. Discussions with George Blumberg of Oxford University were useful in developing ideas underlying section 5. The MOHSST5 dataset was provided by David Parker and Mike Jackson of the Hadley Centre for Climate Prediction and Research. James Murphy, also of the Hadley Centre, provided the data from the UKMO Transient Response Experiment used in figures 18 and 19. Phil Jones kindly provided a copy of his extended Southern Oscillation Index.

Contents

1 Summary

This report presents the scientific case for the development of a 10–12-year data-set of consistent, accurate, global observations of sea surface temperature (SST) from the Along-Track Scanning Radiometer (ATSR) and its successors, the ATSR-2 and AATSR. The case focuses on the role such data will play in detecting and quantifying evidence for anthropogenic climatic change. Four key areas in which data from the ATSR series can play a useful role are identified. They are:

- Providing high spatial and temporal resolution data for the validation of climate models. Reducing the uncertainty in model-based climate predictions is essential before confident, quantitative estimates can be made of the potential social and economic impact of climate change. Many key processes, particularly at the atmosphere-ocean interface, are represented only crudely, if at all, in the current generation of models. Improving the representation of these processes in the models will require consistent, accurate, high-resolution observations of the present climate. Since many processes vary from year to year, multi-year data-sets are essential. The resolution of the ATSR data allows it to be used in the direct validation of the latest generation of eddy-resolving ocean models.

- Providing independent corroboration of *in situ* observations of SST. Current surface-based observations of SST rely on a sparse network of research vessels and drifting and tethered buoys, together with larger numbers of non-specialist "ships-of-opportunity". If decisions are to be based on data from these sources, errors must be quantified as accurately and objectively as possible, to ensure that a warming trend, if detected, is not an artefact of the observing system. Quantifying errors in ship-of-opportunity data is particularly difficult, since observations are made by non-specialists using a wide variety of techniques. Independent satellite observations of ocean skin temperature, together with an appropriate model to relate this to the bulk temperature measured by ships, are needed urgently to check for spurious drifts due to decade-time-scale changes in ship design or the pattern of trading routes, particularly in poorly covered regions. The Advanced Very High Resolution Radiometer (AVHRR) instrument has been shown to be useful for climate purposes only when "anchored" to *in situ* observations (essentially filling in the gaps between them). Thus it cannot be used to provide an independent check in this way.

- Providing direct evidence of global changes taking place. Without knowing precisely what form anthropogenic climate change will take, and at what speed it will occur, we cannot state categorically the probability of it being detected directly by the ATSR series of instruments. If, however, we make some relatively conservative assumptions about natural climate variability on interannual time-scales, and assume the current "best guess" for the rate of warming, we can say

that there is reasonable (>80%) chance of our being able to reject the hypothesis of no secular change at the 97.5% (~2 standard deviations) confidence level after 10–12 years of consistent observations of the global mean sea-surface temperature. This confidence is improved if we model and remove the effect of ENSO. These assumptions also imply that 10–12 years of data provide a >80% chance of rejecting the hypothesis that the warming rate is at the lower end of the current range of uncertainty indicated by climate models if it is, in fact, at the higher end and *vice versa*.

These statements assume that we can be confident of negligible drift in the observing system. The method of on-board calibration of the ATSR makes the instrument much more stable than the AVHRR and less dependent of complex, fallible corrections for non-linearity. This reduces the probability of instrumental drift, and also increases our confidence that, should such a drift occur, we would be able to detect, characterise and quantify it. The dual-angle view significantly reduces the impact of variations in atmospheric parameters such as water vapour and volcanic aerosol on retrieved SSTs. With appropriate surface validation data, this should allow us to discriminate between a secular change in SST, a secular change in these atmospheric parameters (some of which may themselves prove important indicators of global change) and a secular change in instrument characteristics. Identifying the origin of a trend in this way is much more complicated and uncertain with AVHRR data. Likewise, without direct knowledge of the sources of error in ship-of-opportunity observations, we cannot quantify our confidence in them in the way we can quantify our confidence in ATSR, and so exclusive reliance on ship-of-opportunity observations carries an additional risk of some unknown source of spurious drift going undetected.

- Allowing the detection of global patterns of change extending into the Southern Hemisphere extra-tropical regions. An important component of the spatial "fingerprint" of anthropogenic climate change, as predicted by the current generation of models, is a north-south asymmetry, with warming being suppressed in the high-latitude Southern Ocean. Additional structure in the warming pattern near the edges of ice-shelves is also a feature of some models. Detecting the "correct" (model-predicted) *pattern* of change, in addition to detecting a warming trend in the global mean temperature, would considerably enhance our confidence that any global warming is *attributable* to the enhanced greenhouse effect, and not the consequence of some other effect such as a natural climate fluctuation or drift in the observing system (provided, of course, that the patterns associated with these fluctuations were different from that associated with greenhouse warming).

  Detecting global patterns of change requires consistent coverage with the same observing system covering the entire globe. For example, a key component of the predicted pattern is that land areas will warm faster than oceans, but the fact that land surface temperatures are measured in a different way to SST complicates detection of this component of the pattern. ATSR is the only available SST-observing instrument providing global coverage which is considered stable enough for climate change detection purposes. Current *in situ* observations alone are considered inadequate even to define seasonal mean SSTs on a $2° \times 2°$ spatial resolution south of 35°S. Use of unsupported AVHRR data for climate purposes is not justified by the instruments' stability and accuracy. Use of blended AVHRR-*in situ* data is, in turn, problematic, since if a north-south assymetry were detected in blended data it might be attributed to the different weight given to satellite data in the Northern and Southern Hemisphere.

2 Background: the role of data in climate modelling

If an anthropogenic change in the global climate is taking place, then it is clearly important that it is detected and quantified as soon as possible. At present, quantitative predictions of the climate's response to greenhouse gas (GHG) emissions rely entirely on computer models [1]. Models alone, however, cannot provide a reliable basis for climate prediction. Because of our inadequate knowledge of key physical processes, many important parameters in these models must be determined by some form of "tuning" procedure, optimising the fit between model output and observational data. The available data are generally inadequate to determine these parameters unambiguously, and the possibility remains that all models may simply have omitted some unknown process which is having an important effect.

Validating models against climate observations is therefore essential to improve our confidence in model-based predictions of climatic change [15]. The usual starting point for the validation process is to optimise the model's representation of the present-day mean climate. The next step is to investigate the model's ability to simulate climatic variations on interannual time-scales, such as the El Niño / Southern Oscillation (ENSO) phenomenon. The global coverage and high spatial and temporal resolution of data provided by Earth-observing satellite instruments like the ATSR mean they have a clear, and undisputed, role to play in both these tasks, discussed in section 4 below.

An improved understanding of the present day climate and interannual climate variability must, ultimately, improve our ability to forecast climatic change. To put it another way, if a climate model is incapable of simulating interannual climate variations such as ENSO, we would be justified in treating long-term climate predictions based upon it with some scepticism. Many basic physical processes governing the climate system apply to all time-scales, so, for example, an improved cloud-parameterisation scheme developed for a numerical weather prediction (NWP) model can be implemented in a model for climate studies. There are, however, practical considerations which have forced climate research to become, to some extent, compartmentalised by time-scale in the past. The models used for long-term climate prediction are seldom identical to those used for studying shorter-term phenomena, being generally of a much lower spatial and temporal resolution. Simple computational constraints mean that key processes have to be handled differently depending on the objective of the model. This situation may be changing. For example, the introduction of the Unified Model at the UKMO means that essentially the same model will be used both for climate studies and for NWP. Such developments

[1] Some attempts have been made to quantify the climate sensitivity to changes in greenhouse gas levels using statistical analyses of paleo-climatic (ice-core) data: for example [23]. All such analysis (that the author knows of) depend to some extent on computer model predictions, and paleo-climatic studies are primarily used for model validation. See [21] for a useful review

should make validation of models against short-term climatic processes (for which the high-resolution data provided by the ATSR is ideal) even more readily applicable to global change research.

By far the most convincing confirmation of climate-model-based predictions of a long-term climatic change would, however, be the detection of such a change taking place. The reason is that a climate model might perform excellently at the simulation and prediction of ENSO while failing to represent some important effect, such as the cooling due to anthropogenic sulphate aerosols [10], which has significant implications for longer-term climate predictions. Hence the interest in what is normally thought of as "global change detection", *viz.* compiling long-term climate records and examining them for evidence of secular changes which might be attributable to the enhanced greenhouse effect. The role of satellite data in this type of climate research is less clear-cut than its use for direct model-validation. Because of its conceptual simplicity, such research tends to attract the most attention outside the scientific community. It is, however, important to keep it in context.

Ultimately, a long-term climatic data-set is a tool for model validation. Results which are reported from "purely observational" statistically-based research still involve fitting data to models. No statistical test is possible without a hypothesis, which amounts to a model of the system under investigation. Often, such models are remarkably simple, such as the assumption that a climatic data record consists of a linear trend underlying a simple stochastic ("noise") process. Inevitably, reducing data to a form to which statistical tests can be applied involves discarding or neglecting much potentially important information: consider the loss of information involved in averaging sea-surface temperatures (SSTs) over the whole globe to give a single number. Results based on simple statistical models which take no account of our prior knowledge of the physical processes governing the climate system will always be of limited value in advancing scientific understanding.

Climate data-sets also have an important role to play in model "initialization": providing realistic initial conditions for a model-based experiment. At present, this role is only indirectly important for global change research, since most climate change experiments tend to begin with a climate model in an approximately steady state which is then perturbed by an externally-imposed change in boundary conditions, such as increasing CO_2 ("climate prediction of the second kind" [7]). As models become more sophisticated and higher-quality forecasts of changes over the next few years are required for the development of adaptation strategies, direct prediction from "present day" initial conditions will become a priority (prediction of the first kind). High spatial and temporal resolution data such as that provided by the ATSR may prove essential if not only the initial state but also the initial trajectory of the climate system represented by the model is to be determined.

3 Quantifying the impact of satellite data on climate research

The previous section reviewed the fundamental roles of climate data in model validation and initialization. In conclusion, the most convincing argument for extending the ATSR data-set with the AATSR instrument is that the ATSR observational record will then be

- long enough for the investigation of the low-frequency climatic phenomena most relevant to global change research.
- of a high enough resolution, both in time and space, to allow it to be used directly, in conjunction with other data sources, for the validation and initialization of models of many important physical processes.

Consider an example, to illustrate how such a 10- to 12-year data-set may be uniquely suited for the investigation of certain classes of important phenomena. Climate models generally indicate that the surface layers of the oceans will warm much faster than water immediately below them. Suppose, purely as an illustration, that this change in the mean vertical temperature profile were to cause a global change in oceanic eddy activity, which in turn (because eddies may play an important, but difficult to quantify, role in ocean heat transport) were to influence the pattern of global warming. None of the present or immediately forseeable generations of climate models would be able to predict such an effect, since their ocean components are not eddy-resolving. Nor would global SST observations based on ship or buoy data allow the detection of such a change taking place, since they do not distinguish eddy activity. Measurements of sea-surface height based on satellite altimetry provide an indication of overall eddy activity, but if these are to be translated into estimates of eddy heat content, they will, ultimately, need to be supported with eddy-resolving models and high-resolution observations of ocean surface temperature. The only instrument capable of observing SST on this level of resolution, and stable enough for the detection of long-term changes, would be a self-calibrating imaging radiometer like the ATSR.

This is only one of a number of scenarios of difficult-to-detect climatic changes which may occur over the coming decades. It is intended simply to illustrate that there might be climatic changes taking place which we would only be able to detect with an instument like the ATSR, and these changes may have important implications for global change. Because such high-resolution SST data have never been available until now, and the eddy-resolving ocean models required to interpret it are only just being developed, it is impossible to predict what developments we should expect.

Inevitably, therefore, any attempt to quantify the impact of ATSR data on the basis of our present knowledge is likely to underestimate substantially the data's true value. We simply do not know enough about the climate system on the short space- and time-scales resolved by the ATSR to predict the quantitative impact of ATSR data at the scales where it is, ultimately, likely to prove most useful. In response to the Department of the Environment (DoE)'s requirement for a quantitative assessment, this report attempts some very crude investigations, which involve implicitly discarding much of the information potentially contained in ATSR data. In section 5 below, we investigate the potential impact of a 12-year data-set on the detection of a secular trend in the global monthly mean SST. Given the data coverage achieved by the ATSR to date, this means implicitly averaging over 2 million 0.5° longitude by 0.5° latitude "ASST" observations, which themselves consist of an average of up to 2,500 pixels at the full instrument resolution, to give a single number, the monthly mean temperature. We sincerely hope that the ATSR data will be exploited more intelligently than this, but since the ability of an instrument to detect a change in the global mean temperature appears to have become some sort of benchmark of its relevance to global change research, we feel obliged to present these results.

In section 7, we make a very preliminary investigation of the impact of spatial observational coverage on global change detection, with the eventual objective of quantifying the impact of the additional coverage provided by polar-orbiting satellites over that provided by ship and buoy observations. Again, since we do not know enough about the evolution of spatial patterns of SST on long time-scales to use real data, we use output from a climate model as a test data set. The model has a spatial resolution of $3.75° \times 2.5°$. If the real ocean surface temperature were to vary as smoothly as the ocean in such a model, we would not need a high-resolution radiometer to observe it. Thus again, this investigation is likely to understate the true value of the ATSR data.

Before proceeding to these "quantitative" analyses, the following section aims to illustrate the potential impact of the space- and time-resolution of the ATSR, ATSR-2 and AATSR instruments for global change research, using examples from ongoing scientific investigations involving data from the present instrument.

4 The use of high-resolution satellite SST data for the validation of climate models

The utility of satellite SST observations for meso-scale oceanography has been accepted for some time [27]. The objective of this section is to illustrate that the high absolute accuracy, spatial and temporal resolution provided by the ATSR data make it also directly applicable to research into processes and phenomena which are relevant to climate research relating to much longer time-scales. The instrument-validation phase of the first ATSR mission is drawing to a close, and overall results indicate that the instrument is performing to within its design specifications, delivering, in particular, global SST observations with a $0.5° \times 0.5°$ spatial resolution and an accuracy better than 0.5K. This report will not dwell on instrument validation (key results are summarised in Mutlow *et al.* [26]), but it is useful to quote one example, reproduced from Barton *et al.* [5], which serves to illustrate several important points.

Figure 1 shows a particular transect in the Coral Sea in which data from both the ATSR (dotted line) and from the Multi-Channel SST (MCSST) product derived from the Advanced Very High Resolution Radiometer (AVHRR) instrument (solid line) were available, as were both "skin" and "bulk" temperature observations taken by the observing ship. Essentially, the skin temperature is the temperature of the top few microns of the ocean surface, which is what is observed by a satellite-borne radiometer like the ATSR and also, in this case, by a ship-mounted radiometer (asterisks in the figure). The bulk temperature is the mean temperature of the top few meters of the water column. It is observed from bucket measurements, engine-intake observations, drifting and tethered buoys, and, most accurately, a ship-borne thermosalinograph (crosses in the figure).

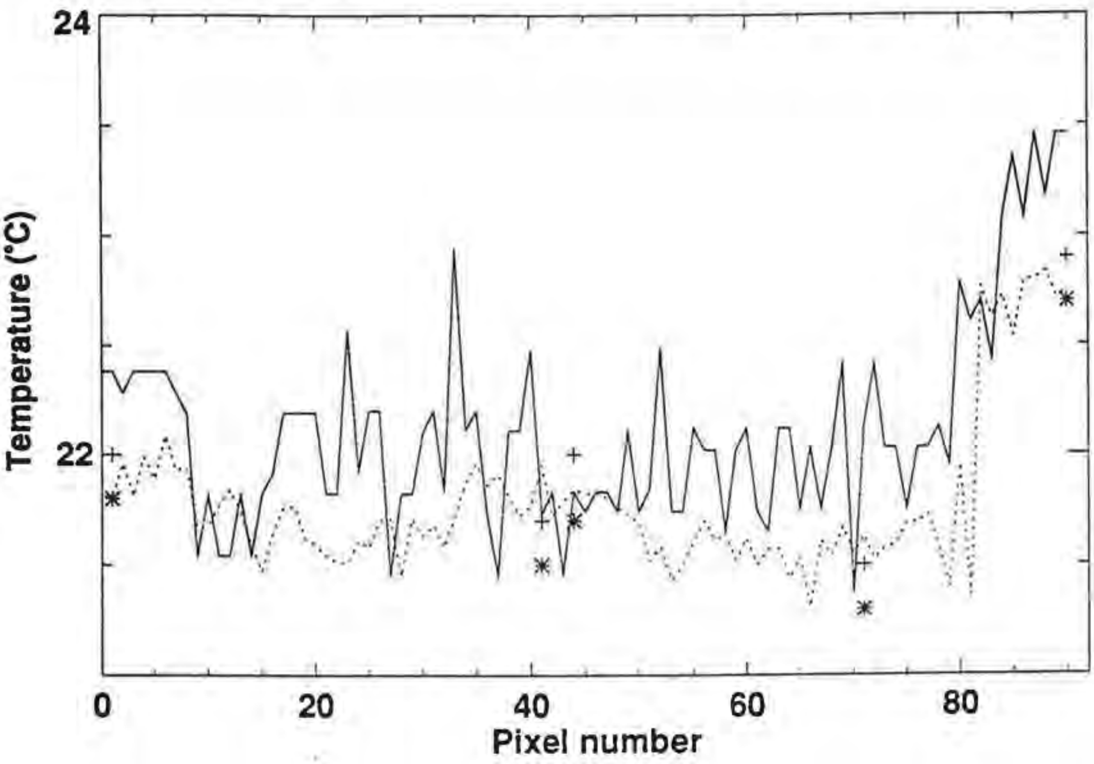

Figure 1: Ship and satellite measurements of SST. Solid line: AVHRR data. Dotted line: ATSR data. Crosses: skin temperature measurements made by a ship-mounted radiometer. Asterisks: bulk temperature measurements made by a thermosalinograph.

Figure 1 indicates that the skin temperature is consistently 0.2–0.3K cooler than the bulk temperature. This is consistent with other observations on this and other validation campaigns, and also consistent with an observed 0.2–0.3K bias between ATSR and drifting buoy data found by Mutlow *et al.*, with the ATSR (skin)

SSTs being cooler than the buoy (bulk) data. This suggests that the correction scheme to compensate for the effect of atmospheric IR absorption has been largely successful, and the ATSR is providing essentially unbiased observations of ocean skin temperature (Barton *et al.* find no significant bias between the ATSR data and observations made by their ship-mounted radiometers).

The ATSR is the first satellite instrument to attempt to deliver an absolute measurement of radiometric skin SST. The AVHRR series provide a set of brightness temperatures which are converted into a pseudo-bulk temperature, the MCSST product, using a completely empirical regression formula based on buoy observations. Thus the same algorithm is used for AVHRR both to correct for the skin-bulk temperature difference, and to correct for effect of the atmosphere. Wick *et al.* [46] discuss the deficiencies of this scheme. The AVHRR observations are inherently less reliable than ATSR, since they include, with the instrument variance, the variance of the buoy observations used to develop the empirical model. The correction algorithms can also be subject to regional bias, being dominated by conditions in regions where buoy data exists.

The climate modelling and NWP communities, however, are accustomed to dealing with bulk SST, a factor which may have delayed initial utilisation of ATSR data. Developing physically-based algorithms to derive bulk from skin SST using local wind and sea-state data is currently a priority for the ATSR Science Team at RAL. The skin-bulk temperature difference has also been a persistent issue in validation campaigns.

While providing problems for immediate validation and use, the fact that the ATSR measures the ocean skin temperature may, ultimately, provide an opportunity to investigate an issue which is of crucial importance for global change: atmosphere-ocean exchange of CO_2. Very recently, it has been observed that the skin-bulk temperature difference may play a crucial role in modulating oceanic uptake of CO_2 [34], with the small inter-hemispheric assymetry in the average skin-bulk temperature difference possibly accounting for a large fraction of the well-known "missing sink" of CO_2 in the northern hemisphere [40]. In the study presented in ref. [34], a simple model was used to calculate the skin-bulk temperature difference as a function of local wind speed and solar radiation. The authors were obliged to use monthly-mean climatological data for their calculations, and emphasised that their confidence in their results would be considerably enhanced by higher resolution data. Provided adequate surface-based radiometric skin temperature observations were made to validate the ATSR data for a wide range of locations and atmospheric conditions (also currently a priority activity at RAL), it might be possible to use ATSR data in conjunction with surface-based bulk temperature observations to develop observationally-based global maps of the skin-bulk temperature difference. The importance of such research would extend well beyond atmosphere-ocean CO_2 exchange, since the skin-bulk temperature difference is a key indicator of the vertical heat flux at the atmosphere-ocean interface. Direct observation of this heat flux represents something of a holy grail in studies of atmosphere-ocean interactions.

Although the short length of the present ATSR data-set precludes the direct investigation of interannual time-scale climate phenomena, a number of research projects making use of ATSR data are already proposed or in progress which are of direct relevance to our understanding of processes important for global change. They include:

- using ATSR data to validate a near-eddy-resolving general circulation model (GCM) of the Mediterranean Sea for studies of deep-water formation (with the University of Edinburgh). Deep-water formation, the process by which surface waters, containing dissolved anthropogenic GHGs and (potentially) anomalous heat due to global warming, penetrate to the ocean depths, is a key factor for the oceans' role in climate change. It takes place in a relatively small number of locations, the most accessible being in the Mediterranean.

- using ATSR data in conjunction with stratospheric water-vapour observations from the Microwave Limb Sounder on the Upper Atmosphere Research Satellite to investigate stratosphere-troposphere exchange of water vapour (with the University of Reading). Stratospheric water vapour is a very potent greenhouse gas, so an improved understanding of its sources and sinks is clearly important for climate research. The dual-view configuration of the ATSR instrument may improve our ability to observe penetrative convection events in the tropics, which play an important role here.

- using ATSR data to study the optical properties of thin cirrus cloud in the water vapour window region of the spectrum. The ATSR channels were specifically selected to lie in the region of the infra-red (IR) spectrum where absorption by atmospheric water vapour is weakest. For the same reason, this may be thought of as one of the few spectral regions through which the Earth's surface sheds energy directly to space. Quantifying the impact of clouds on the Earth's radiation budget is a priority for global change research. Thin, high cirrus clouds pose particular problems, because they are difficult to identify with the visible frequencies used by meteorological satellites to quantify cloud cover. Prata and Barton [28] have observed that the information contained in the different radiance channels of the ATSR, in conjunction with the high-quality SST retrievals possible with dual-view observation of the surface, may be used to study the optical properties of such clouds. For example, on the upper right-hand corner of figure 2 we observe an ocean front near Japan with an overlying thin cloud (the white band running across the corner). In this situation, we have an accurate estimate of SST under the cloud, inferred from SSTs immediately adjacent to it, and there is also a strong SST gradient under the cloud. Radiation from the surface is partially transmitted through the cloud (the front is clearly visible in this image), allowing the way the cloud absorbs IR radiation to be studied in detail.

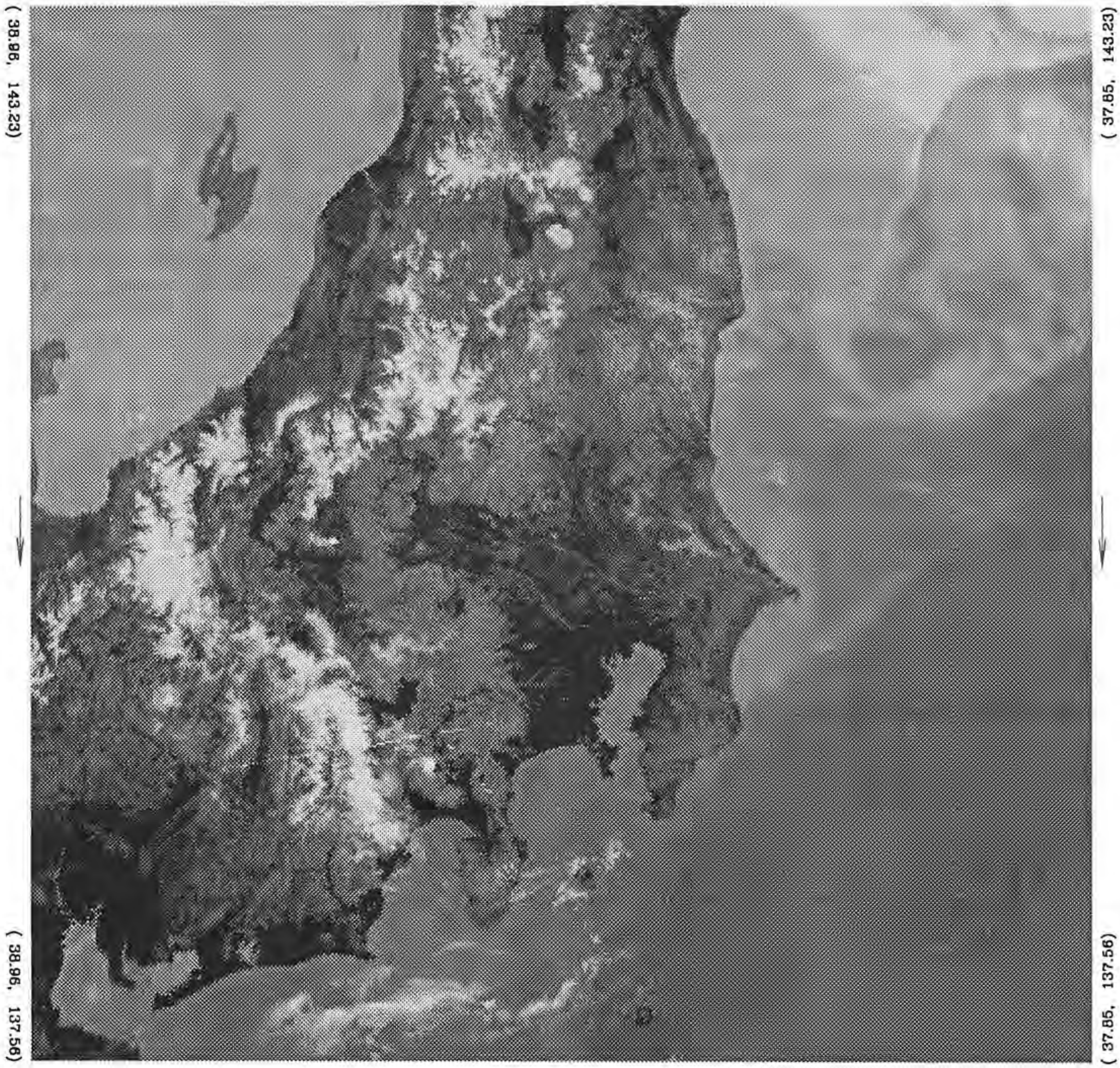

Figure 2: ATSR 11μm nadir image of Japan. Snow-covered mountains appear white (radiatively cold) in this figure. Note thin cirrus cloud (cold, and thus white, in this figure) overlying a strong ocean front in the upper right hand corner.

- using ATSR data in conjunction with a medium-resolution ocean GCM of the Pacific basin for studies of equatorial atmosphere-ocean interactions (with the University of Oxford). Such interactions play a crucial role in interannual climate variability, particularly ENSO. As was argued above, an improved understanding of the behaviour of the climate on these time-scales must improve our confidence in longer term climate prediction.

This last item provides an excellent example of the type of model-data interaction which, this report has argued, represents the best use of the ATSR data. The following results are unpublished, and are presented here on the understanding that this report will not be widely circulated. Figure 3 shows a time-longitude section ("Hovmoller diagram") across the Pacific basin of SST generated by a medium resolution ocean GCM forced with daily wind-stress data from the European Centre for Medium Range Weather Forecasting (ECMWF) analyses [22]. The horizontal scale represents longitude from 120°E to 60°W, and the vertical scale represents time, increasing upwards. The period shown is 1992, and the latitude is 2.25° north. The colour scale, red-yellow-green-blue, represents warm to cold SSTs respectively. The annual seasonal cycle, modulated by the 1992 El Niño event, is clearly visible in the East Pacific (the right hand side of the image). In addition, we observe remarkably regular waves propagating from east to west near the centre of the image.

Similar waves have been observed in equatorial models before, and were commonly assumed to be due to shear instabilities on the edge of the equatorial current ("Legeckis waves"). If this were the case, however, we would expect such waves to be initiated at random, with random phases. Figure 4 shows ATSR combined day- and night-time ASST ($0.5° \times 0.5°$ spatial resolution) data, interpolated to the model grid, for the same period [3]. The regular pattern of gaps in the data is due to the 35-day repeat cycle of the ERS-1 satellite through most of this period. Remarkably similar wave activity also appears, at the same time and in the same region.

The fact that these waves appear at precisely the same time and have the same propagation velocity in both data and model suggests that they are not simply an instability, but are forced by, or at least phase-locked to, some periodic signal in the wind-stress. They appeared much weaker, if at all, in 1991, indicating they are probably related to the 1992 El Niño event. The very slow propagation velocity of these waves means they would take over a year to traverse the basin. Slowly-propagating equatorial waves almost certainly play an important role in ENSO, although the significance of these particular waves remains to be investigated.

These waves were originally identified by the application of a novel signal-processing technique, Singular Spectral Analysis (SSA) [45] to the satellite data, but we were initially sceptical of their origins because the period is coincidentally very close to that of the orbit repeat cycle. Their appearence in the model, which only receives data on the real world via the ECMWF analysed wind-stress, indicates that their origin is almost certainly geophysical, and neither an artefact of the model numerics nor of the satellite orbit. This is a classic example of the interaction between models and data increasing our confidence in both.

This section has attempted to indicate how the high spatial and temporal resolution of ATSR data may be used to improve our understanding of key climatic

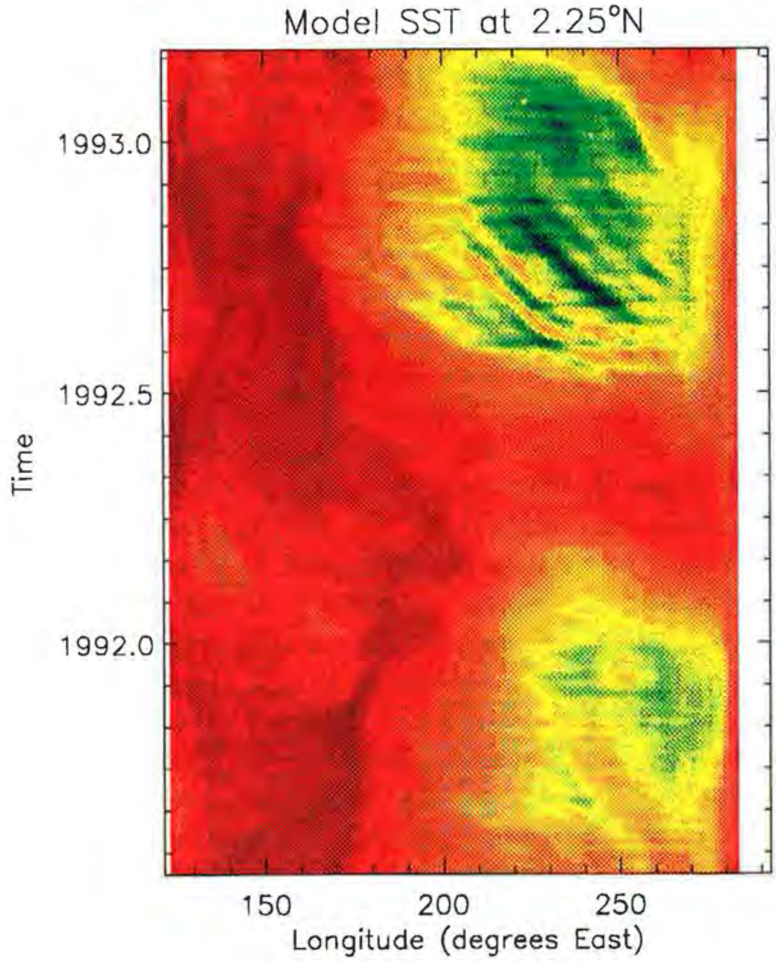

Figure 3: Time-longitude section of SST at 2.25°N across the Pacific basin (from 120°E on the left to 60°W on the right) from a medium-resolution ocean model forced with daily wind-stress analyses. Warm SSTs are red, blue cold. Note regular east-to-west propagating waves above the centre of the image.

processes, and thus improve our confidence in the climate models used for global change prediction. This, in the view of the author of this report, represents the aspect of ATSR data-utilisation which will be of most interest to the scientific community. There is, however, a case to be made for using such data to address much simpler issues, such as attempting to detect a secular trend in the global mean sea-surface temperature. The case rests not so much on the scientific value of such research in furthering our understanding of the climate system (simply detecting a trend in a data-set without any attempt to fit the data to a physically-based model of the underlying system is inevitably a somewhat meaningless activity), but on its value to the wider community in confirming the overall thrust of global change research. If ATSR data is to be used for this purpose, the question needs to be addressed as to whether we can expect to detect any significant trend in a data-set of only 12 years.

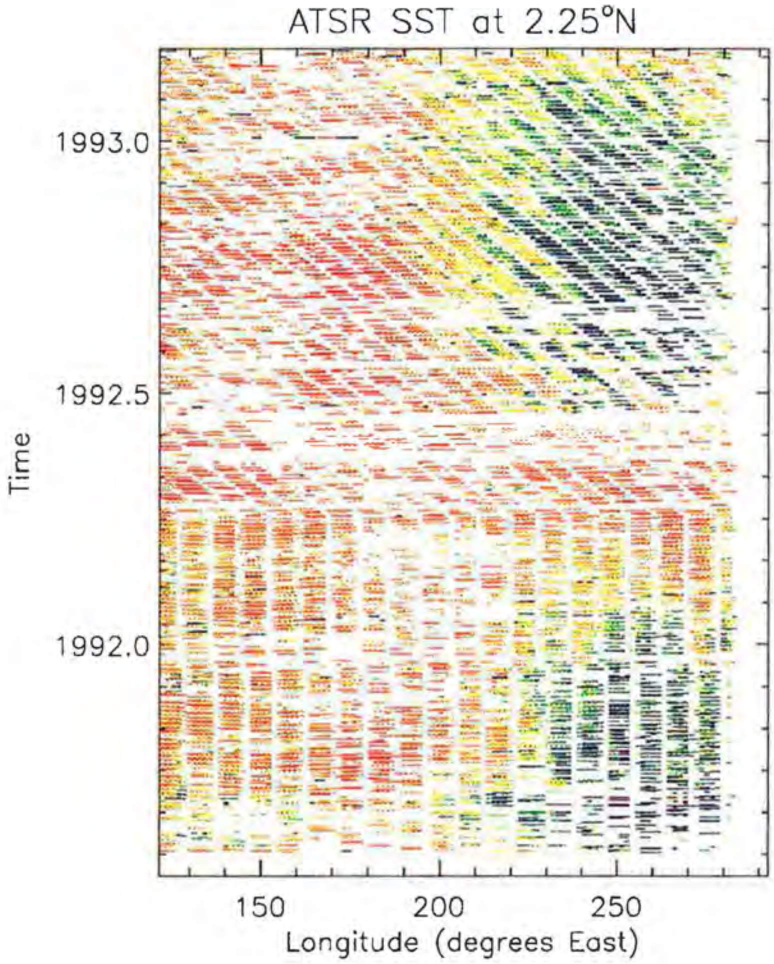

Figure 4: Time-longitude section of SST at 2.25°N across the Pacific basin (from 120°E on the left to 60°W on the right from ATSR combined day-time and night-time ASST data. Colour scale has been adjusted to correct for a 2K warm bias in model, but is otherwise identical to previous figure. East-to-west propagating waves also appear, coherent with waves in previous figure, indicating these are not instabilities, but have a common origin in some form of ocean-atmosphere interaction.

This is considered in the next section.

5 Prospects for global change detection in a 10–12-year dataset

Recently, there has been considerable interest in the detection of a secular trend in global temperatures using newly-emerging satellite data-sets such as those derived from the Microwave Sounding Units (MSU) on the Tiros-N series [38] and MCSST data from the AVHRR instrument [39]. In the case of the MSU data, concern has been expressed that individual climatic events such as the very strong 1982–3 El Niño event and the volcanic eruptions of El Chicón (1982) and Mount Pinatubo (1992) may have introduced significant biases into such short records [13]. Even more serious doubts have been raised about the use of unsupported AVHRR data for climate research (see section 6 below for a discussion of calibration of the AVHRR and ATSR instruments and the problem of stability of AVHRR data).

Issues of data accuracy and stability must be addressed through technical evaluation of specific instruments. A more general issue is whether, given the level of natural variability in the climate system, we should *expect* to be able to detect a global warming trend of the magnitude predicted by climate models in a decade-long data-set. This question is clearly pertinent to the direct application of data from the ATSR series to global change detection. Should we find that the chance of detecting a significant trend in global temperatures in a 10–12-year data-set was virtually nil, then it would be unjustifiable to suggest that the ATSR data would be likely to make a direct contribution to the detection of such a trend.

Even in this situation, the ATSR data-set could still make an important indirect contribution, by providing much-needed corroboration of *in situ* observations, in particular to look for systematic variations due to secular changes in sampling in data-sparse regions. This is an obvious application of the data-set, guaranteed to be of value to global change research. The reason the ATSR instruments provide an independent data-set in this way is that the effect of the atmosphere is taken into account in the analysis of ATSR data using a physically-based model determined prior to the instrument's launch and validation. In contrast, the AVHRR-MCSST product uses an empirical regression formula between satellite-observed brightness temperatures and collocated *in situ* observations to convert the former into an estimate of SST. This formula is revised with each new instrument. Thus long-term trends in AVHRR data will be entirely determined by the low-frequency behaviour of the *in situ* data used in this empirical model, and as such cannot be used as an independent data-set to check the *in situ* record.

For global change detection, however, the benefits of independent validation of the *in situ* record would be indirect: should the ATSR increase our confidence in the accuracy of the ship-based record, then we would clearly be more inclined to believe evidence of global change derived from the ship data. But should we expect to be able to detect a global change in SST based on the ATSR data alone?

For the purposes of this section, we will focus on the global mean sea-surface temperature. Clearly, this is a very crude detection statistic, and (as was remarked above), not a parameter which demands (and thus "shows off") the precision and resolution of the ATSR data-set. But it is, at least, familiar, and if we can show there is a reasonable chance of the ATSR data allowing a significant trend to be detected in this very crude statistic, then the use of a more sophisticated statistic must increase that chance. The IPCC [19] estimated that we should expect to detect an increase in the global mean temperature some-time between 2002 and 2047, depending on the true rate of warming, but this was based on the completely arbitrary criterion that we should not accept a warming as significant until a further 0.5K rise was realised above present-day temperatures. This section investigates whether the expected detection time can be estimated more objectively.

5.1 The problem of inter-decadal oscillations

Detecting evidence of a secular climatic change is complicated by the extent and complexity of natural climate variability on all time-scales. For example, Ghil and Vautard [16] claim to identify an interdecadal oscillation in global temperatures which, they claim, accounted for a significant proportion of the anomalous warmth during the 1980s. Should such an oscillation exist, then whether or not we detect a warming trend in a 10-year data-set would be as much a function of where we are relative to the phase of this oscillation as it would be a function of the magnitude of the underlying trend itself. This remains an area of considerable controversy. Allen and Smith [4] demonstrate that the stationary component of the global annual mean temperature record can be adequately characterised as a simple linear stochastic ("red noise") process, calling into question the observational evidence for interdecadal oscillations. Decomposition of an observational record into secular, oscillatory and stochastic components is particulary problematic near the end-points of the series [1]. Unfortunately, these are generally the areas we are interested in: if we need to know the state of an interdecadal oscillation in order to interpret results from a new satellite data-set, we are far more likely to be interested in the period 1984-93 than the period 1954-63, although statistically, the problem of what the oscillation was doing in the 1950s is much better constrained.

Thus the case for genuine (in the sense of significantly predictable) oscillations in global temperatures on interdecadal timescales remains to be proven. Should such osçillations exist, then no 10–12-year data set, from any instrument, could be correctly interpreted without explicitly taking them into account. If this proves to be the case, then the first priority for the ATSR data-set will be the understanding of natural decadal time-scale climate variability. The hope here would be that, given the high resolution of the ATSR data, it may be possible to develop and validate physically-based models of any decadal and interdecadal oscillations which would allow us to estimate and predict their influence on global temperatures, somewhat as Jones [20] advocates we take into account the effect of ENSO. We would then look for evidence of secular trends in the residual. On the other hand, the most conservative assumption at present seems to be that no genuine interdecadal oscillations exist, and the considerable variability in global temperatures on such time-scales represents nothing more than auto-correlated ("red") noise.

5.2 Using past observations of global mean SST to estimate probable detection times for global change under different scenarios

We can obtain an approximate upper limit on the estimated natural variability of the global monthly mean SST by using data from past observational records and assuming that the stationary variability observed in the past results from genuine natural climate fluctuations which will continue in the future. This is only an approximate upper limit, since the observational record could either overestimate or underestimate the magnitude of the natural variability. It probably overestimates it, since the record will also include noise due to measurement error and variations in sampling coverage. Increasing coverage will not invariably reduce the level of stationary variability, since it is always possible that newly-covered regions may be strong sources of interannual climate variability.

The question we wish to address in this section is as follows: *if* the stationary component of the natural variability of the global temperature remains as observed over the past 90 years, *and* there really is a linear underlying warming trend, then what is the probability of our detecting this trend at a given confidence level in a 12-year data-set? This question is sufficiently obvious that it must have been addressed at some stage, but being unable to locate a reference, we are obliged to present some simple calculations here.

Figure 5 (dotted line) shows the monthly mean SST compiled by taking a global average of $5° \times 5°$ monthly anomalies about the 1951-80 climatology, where-ever data was available, weighting by the cosine of latitude to take into account the variation of the area of grid-squares. The data is taken from the UKMO MOHSST5 data-set [8], kindly provided by M. Jackson and D. A. Parker. We have not corrected for coastal grid-squares being partially covered by land, which would introduce a bias towards coastal observations, but comparing seasonal anomalies calculated in this way with those shown in ref. [8] indicates that the overall effect of this approximation is quite small.

All variability on timescales of >40 years has been extracted using a sliding window technique, based on SSA, which allows for the presence of a non-linear trend, adapted for unbiased trend-reconstruction near the series end-points as advocated in ref. [2]: see appendix A for details. The window width used is 480 months (40 years). We also extract all variability on the period of the annual cycle. This is already low, since these data are expressed as anomalies about the 1951-80 climatological annual cycle, but it cannot be assumed to be zero throughout the series, since the geographical pattern of coverage was very different in the last century to the pattern today. Nor would we be justified in calculating a single annual cycle for the complete series, since its amplitude (and even phase) may vary over time due to the changing coverage. The combined reconstructed trend and annual cycle are shown in figure 5 (solid line). The appearence of a small annual cycle during 1951-80 is simply a consequence of

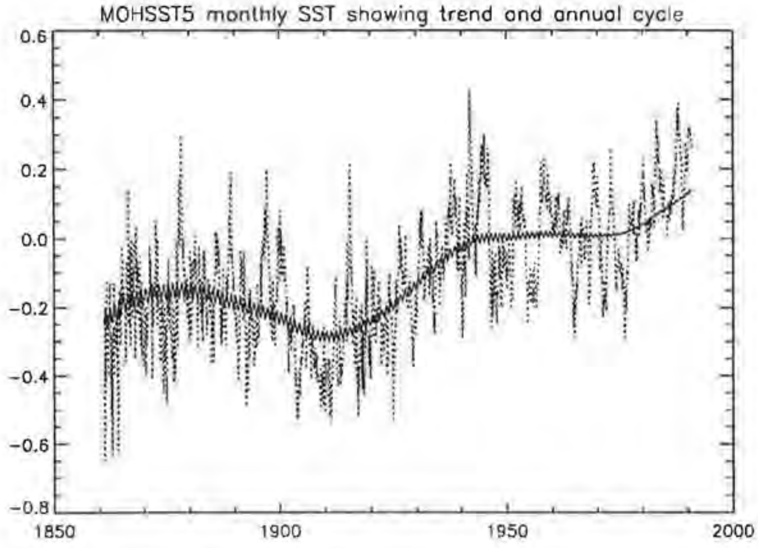

Figure 5: Dotted line: Monthly values of global area-mean SST for the years 1861–1990, averaging anomalies about the 1951–80 climatology from the UKMO MOHSST5 data-set. Solid line: reconstructed trend and annual cycle using a 480 month sliding window.

the sliding window technique. The annual cycle in the years immediately before and after 1951-80 will have "contaminated" the reconstruction during that period.

The low-frequency component of the reconstruction shows some evidence of a warming trend, with some 50–100 year time-scale variability superimposed upon it. Analysis of this record alone cannot tell us whether this apparent trend represents a secular change or if it is simply one component of a multi-century time-scale oscillation. Nor can it tell us whether the 50–100 year activity represents oscillations, random fluctuations or a deterministic aperiodic process. The record is simply too short. Climate models used to simulate the response to an exponential increase in greenhouse gas concentrations (approximately what is occuring at present) tend to indicate a warming which is close to linear on >40-year time-scales: the radiative forcing change due to an increase in CO_2, in particular, is proportional to the logarithm of the perturbation CO_2 concentration [17]. If all the changes in figure 5 represent real climate variations, there are (unsurprisingly) processes taking place on >40-year time-scales which are currently not well represented by the models. Investigating these processes requires physically-based models, validated against much longer data-sets such as climate reconstructions over the past few thousand years from tree-ring and ice-core records.

If these low-frequency phenomena were to compensate for the warming trend over the coming decade, as they appear to have compensated for whatever warming trend there is in figure 5 during the period 1940–75, then no observing system, however sophisticated, would be able to detect a trend in the global mean SST over the next few years. This does not necessarily mean that we would be unable to detect an anthropogenic climatic change, because the global mean SST is by no means the only detection statistic available. In a sense, we have already detected an anthropogenic climatic change, in the form of a steadily rising global mean radiative forcing due to long-lived GHGs, inferred from GHG concentrations. The following section will discuss one of the alternatives to simply focussing on the global mean temperature. The detection of climate change is much more subtle problem than a simple question of "is it warming up yet?" [47]. The use of more sophisticated detection statistics requires precisely the sort of global, high-resolution data provided by the ATSR. However, should we fail to detect a rise in the global mean temperature (or should we place an upper limit on the rate of warming which is lower than that predicted by some climate models), this would at least indicate that those models which predict a warming larger than that found in the data are omitting something important.

Our concern here is not to evaluate the significance of the warming trend in figure 5 itself, but to assess the implications of the stationary variability in this series, which we assume to correspond to the variability on <40-year timescales, for the detection of such a trend were one to exist. To do so, we will adopt a simple, relatively optimistic scenario for the components of variability in figure 5 concerning which we have no quantitative information: we will assume that, from now on, all the >40-year variability in global temperatures will conform to a linear trend with a gradient which we will specify. The statistics of the <40-year component of variability in figure 5 are, as far as we can tell, stationary for the period 1901 onwards. The variance is somewhat higher before that time, presumably due to poorer coverage (the accuracy of individual ship-of-opportunity observations has apparently changed very little: P. Jones *pers. com.*).

We therefore construct a set of 11 artificial 90 year (1080 month) timeseries, using the following model

$$T(t) = \beta(t - t_0) + S(t) \tag{1}$$

where $S(t)$ is the stationary component of the observed global monthly mean SST with the annual cycle removed (the dotted line minus the solid line in figure 5), t is the time, with t_0 arbitrarily set to January 1946. Thus we have replaced all the variability shown by the solid line in figure 5 with a linear warming trend, the $\beta(t-t_0)$ term.

β is an externally specified parameter taking one of the values 0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.045 and 0.05 in each of the 11 series. These span the range of warming rates in degrees per year proposed by the IPCC [24] on the basis of model simulations of the climate's response to increasing GHG levels. The IPCC's "best estimate" of the warming rate over the coming century is 0.025°C per year, with a confidence range of 0.015 to 0.04°C per year, depending on the climate sensitivity (see figure Ax.2 in ref. [24]). Whether we should expect this rate of warming to be realised during the 1990s depends on how fast the warming "takes off" in response to GHG emissions to date. Integrations of coupled models forced with the observed pattern of increase of GHGs over the past century are still in progress (Stouffer, *pers. comm.*). 0.025°C per year seems a reasonable estimate of the expected warming through the 1990s given our current knowledge [24].

We now investigate the number of years of data required, on average, to detect this trend in each of our 90-year test data-sets. Our interest in this issue arises from the fact that the ATSR series of instruments will generate a 12 year SST data-set which, it is claimed, will be applicable to the problem of global warming detection. Should we find that, even if there is a perfectly genuine linear warming trend in these artificial time-series, the number of years required to detect it is considerably greater than 12, then it would be possible to question the relevance of the ATSR data-set to this problem. As we will see, this is not the case.

Detecting anything invariably means ruling out a null hypothesis at a prescribed confidence level. The present section shall focus on the hypothesis that there is no warming trend in the global mean SST: i.e. that the climate sensitivity to increased GHGs is effectively zero. Since no realistic climate model suggests this is the case, this may appear a rather pointless hypothesis to consider. It is, however, the hypothesis almost invariably adopted in "purely observational" studies of global change detection, so we are at least following general practice. We shall look at a couple of more realistic hypotheses in conclusion.

The number of years of data required for trend-detection depends on the size of the trend in the detection statistic and on the properties of the stationary variability, or "noise", represented here by $S(t)$. In addition, the probability of detecting a trend in a given number of years of data depends on the level of risk which we are prepared to accept of a false-positive result: *viz.* detecting a spurious trend in data where there is none. This is specified by the level of significance demanded in a statistical test. Clearly, if we are cautious, demanding a very low probability of a false-positive, then we also increase the amount of data required for the detection of a trend of a given amplitude, or conversely, increase the risk of our failing to detect a genuine trend in a given length of data. We adopt a confidence level of 97.5% throughout, which corresponds to approximately two standard deviations for a one-tailed t-test (we are only attempting to detect a positive trend: hence the test is one-tailed).

Ordinary least squares regression is inappropriate for data of this nature, since the noise component $S(t)$ is heavily auto-correlated in time (1-month lag autocorrelation of 0.84). Instead, we use estimated generalised least squares (EGLS) regression [11] with a first-order auto-regressive, or AR(1), noise model. Maximum-likelihood estimators give similar results, but were found to be subject to a larger bias when tested on artificial data.

The monthly mean global SST series has not yet been investigated in detail to establish the adequacy of an AR(1) noise model as a representation of its detrended component. However, a detailed analysis of the annual mean series [4] indicates that the AR(1) model is appropriate in that case, and subsequent analysis along the lines pursued in this section using annual-mean data give very similar results to those quoted here.

We test overlapping short segments of the artificial series $T(t)$ for a positive linear trend, using EGLS regression with time as the independent and $T(t)$ as the dependent variable. Figure 6 shows how the probability of detecting a significant trend (i.e. rejecting the null hypothesis of no trend at the 97.5% significance level) in a segment of a give length varies with β, the magnitude of the actual trend in $T(t)$. Probabilities have been estimated simply by calculating the proportion of trial segments to give a positive result. For all trend and segment-length combinations lying above the solid line in figure 6 there is a $\geq$50% probability of detecting a significant positive trend in a selected short segment of $T(t)$. Likewise, for all trend and segment-length combinations lying above the dashed (dotted) lines we have a $\geq$80% ($\geq$95%) probability of detecting a significant positive trend.

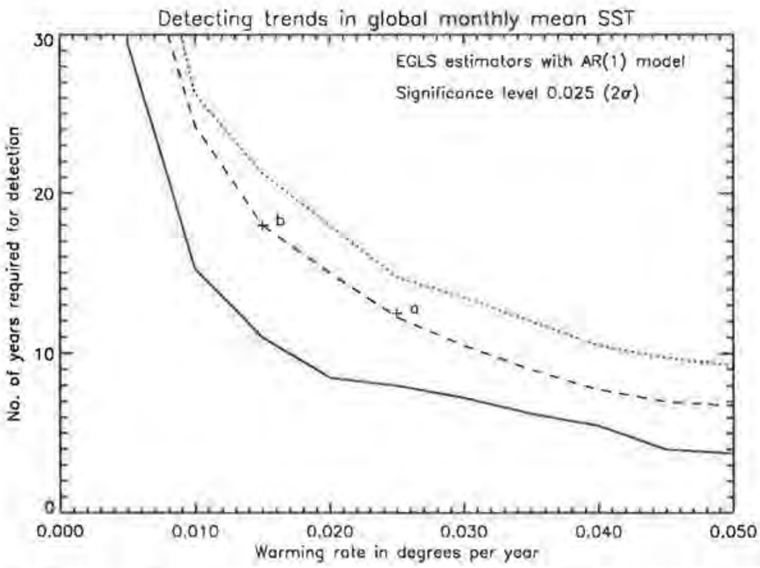

Figure 6: **Number of years required for EGLS regression to indicate a significant positive trend in a short segment of $T(t) = \beta t + S(t)$ ($S(t)$ is the detrended, deannualised monthly mean SST from 1901 to 1990 and t represents time), as a function of the actual magnitude of the trend, β, in 50% of trials (solid line), 80% of trials (dashed line) and 95% of trials (dotted line). Detection implies rejecting $\mathcal{H}(\beta \leq 0)$ at the 97.5% confidence level. 12.5 years of data are required for an 80% probability of detection if the trend is only 0.025°C per year (point a), and 18 years are required for an 80% chance if the trend is only 0.015°C per year (point b).**

Point a in figure 6 indicates that, with the 12.5-year data-set provided by the ATSR series, we have a >80% chance of rejecting (at the 97.5% confidence level) the hypothesis of no trend in $T(t)$ if the background trend is 0.025°C per year (the "best guess" proposed by the IPCC). Point b indicates that we require 18 years to have a >80% chance of rejecting $\mathcal{H}(\beta = 0)$ if the background trend is 0.015°C per year (the IPCC's "low" estimate).

These probabilities are not small, and indicate that we do indeed have a relatively high chance of detecting a global change in SST in a 12 year data-set if the predictions of current climate models are correct and we assume a perfectly stable observing system. On the other hand, detection of a perfectly genuine trend is far from guaranteed, particularly if the gradient is at the weaker end of the uncertainty range. Much of the popular discussion about the implications of the absence of any significant trend in MSU data over the 1980s is therefore somewhat misleading.

We have been optimistic in our assumptions concerning the trend itself, in assuming it to be perfectly linear, and therefore ideally suited to detection using linear regression-based algorithms. On the other hand, we have also been very conservative in the way we have handled the stationary variability of the series, assuming we know nothing more about it than that it more-or-less conforms to an AR(1) noise model. This is not, in fact, the case. For example, we know that a proportion of the interannual variability in global temperatures can be accounted for by the ENSO phenomenon. An effective indicator for the state of ENSO is the Southern Oscillation Index (SOI), defined, in the version shown here following ref. [20], as the normalised difference in sea-level pressure between Tahiti and Darwin, with the individual station data themselves normalised and centered about their respective 1951-80 climatological annual cycles. The dotted line in figure 7 shows the monthly SOI extended back to 1870: data kindly provided by P. D. Jones of the Climate Research Unit of East Anglia, and the solid line shows the annual cycle and >40-year variability extracted as above.

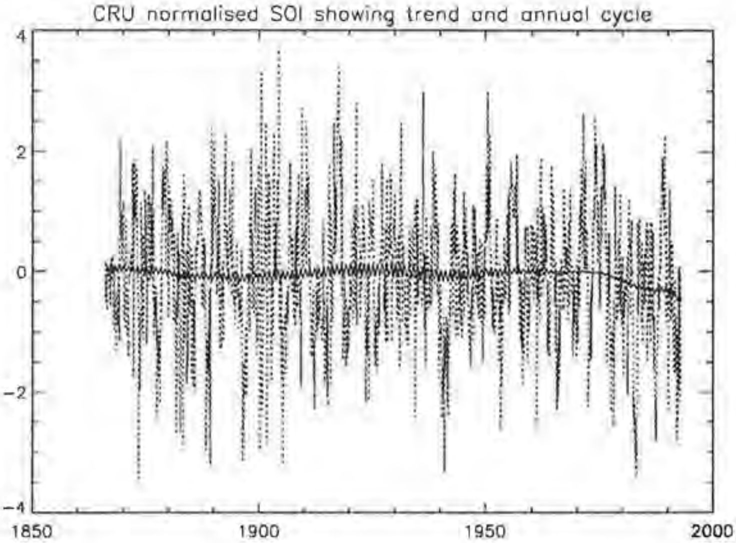

Figure 7: Monthly mean normalised Southern Oscillation Index 1870-1992, normalised and centered about the 1951-80 mean annual cycle (dotted line), and residual annual cycle and trend extracted with a 480-month running window (solid line). Data provided by P. D. Jones of the CRU, East Anglia.

The relationship between the SOI and the global temperature is clearly shown by figure 8, which shows the product-moment correlation between the two series after their respective trends and annual cycles have been removed. No further time-domain filtering has been applied. We see a remarkably clear peak of negative correlation with the SOI leading global SST by 4–5 months. The positive peak occurring approximately 20 months earlier is a consequence of the cyclic nature of ENSO (its dominant period is of the order of 40 months).

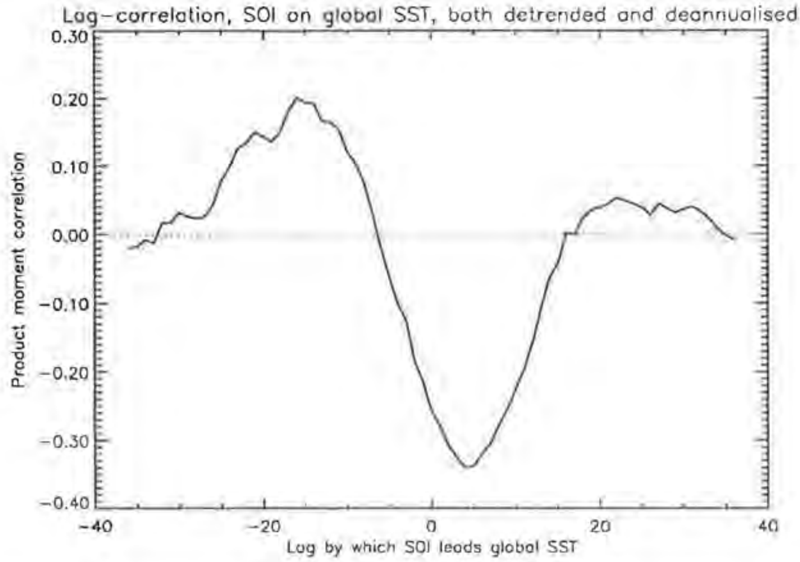

Figure 8: **Product moment correlation between the normalised SOI and the global SST series, both with trends and residual annual cycles removed, as a function of lag in months by which the SOI leads the global SST. A negative SOI, corresponding to a warm El Niño event, is a precursor of anomalously warm global SSTs at a lead of 4–5 months.**

Suspecting that this relationship between ENSO and the global temperature might only apply to the latter part of the time-series, since coverage of the Pacific was so poor in the last century, we apply a sliding window based regression technique to the two series (see appendix), to see if there is a secular trend in the magnitude and significance of the coefficient relating them. Results are shown in figure 9. We are interested in explaining variability in global SST in terms of the SOI, so we take the SOI as the independent variable in the regression and SST as the dependent variable. For this reason, it is necessary to noise-filter the SOI by applying a 9-month running mean prior to the regression. With only a short data set, it would not be possible to filter the SOI in this way, thus the method described here should not be taken as a practical method for eliminating ENSO-related variability from SST data-sets, although after further work, it might form one component of such an activity.

If the relationship between ENSO and the global temperatures really did emerge only in data from this century, we would expect to see a clear trend in the regression coefficient (solid line), and also expect the 2σ limits (dotted lines) to lie on either side of the zero line in the last century. Neither expectation is borne out by figure 9: ENSO is clearly a sufficiently global phenomenon that it appeared in the global temperature records of the last century, even though the coverage of the Pacific was

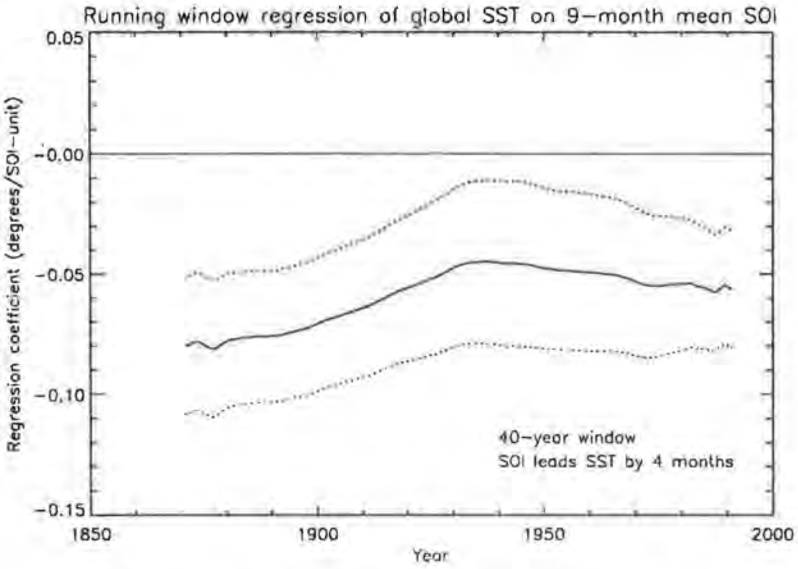

Figure 9: Variation of a running window regression coefficient between the global monthly mean SST and a 9-month running mean SOI, with the SOI leading the SST series by 4 months. Solid line shows regression coefficient, dotted lines show upper and lower 2σ limits, all averaged over a 40-year window. Coefficients are significant and negative throughout the period considered, and there is no evidence of a significant secular trend in the relationship between SST and ENSO.

so poor.

We reconstruct the component of the global SST which is "explained", in a statistical sense, by the SOI, using this regression model with time-varying coefficients, and subtract it from the SST series along with the trend and residual annual cycle. We then repeat the analysis applied to the original detrended, deannualised SST series to see if removing the variability attributable to the SOI has improved our ability to detect secular trends in global temperatures. Figure 10 shows the result. There is some improvement: we now only need an 11-year data-set to have an 80% chance of detecting a positive trend given the IPCC's "best" estimate of the rate of warming (point a).

This figure is simply intended to illustrate how an understanding of interannual climate variability may improve our prospects for global warming detection. The improvement in detection time between figures 6 and 10 will definitely contain some level of "artificial skill", since the same data-sets used for developing the statistical model of the relationship between the SOI and the global SST were also used for testing the impact of the model on global change detection. Rather than relying on purely statistical methods, as here, it would clearly be preferable to develop a physical understanding of the mechanisms relating ENSO to global temperature and model its impact explicitly. This is one example of the benefits which accrue to global change research from reseach into shorter time-scale climate variability. This analysis is only intended to indicate what might be achieved by such an exercise.

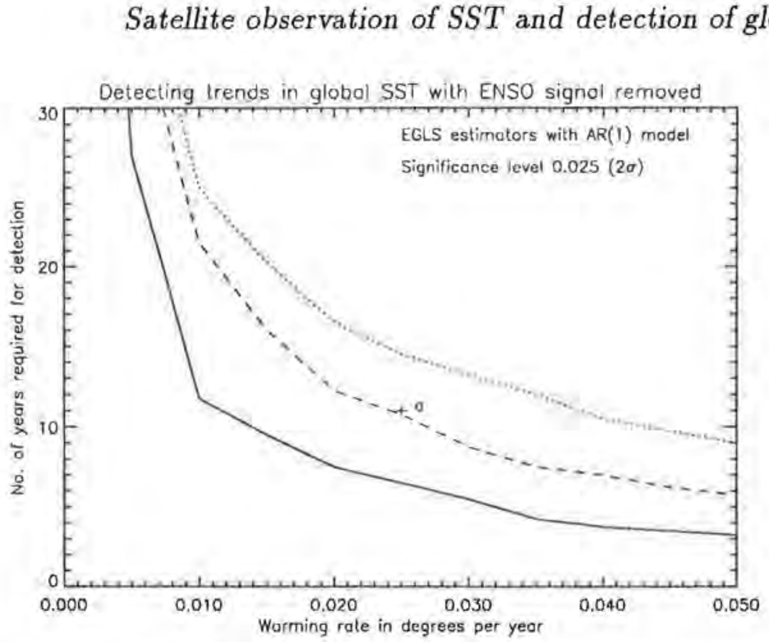

Figure 10: Number of years required for EGLS regression to indicate a significant positive trend in a short segment of $T(t) = \beta t + S_2(t)$, where $S_2(t)$ is the detrended, deannualised monthly mean SST from 1901 to 1990 with the component explained by a best fit on a 40-year window to a 9-month running mean SOI also removed. Dotted, dashed and solid lines as in figure 6 above. Only an 11-year data-set is required for a >80% chance of detection if the trend is 0.025°C per year (point *a*).

Thus far, we have focussed on rejection of the null hypothesis of no trend in global temperatures (zero climate sensitivity). As was remarked above, this is not necessarily a particularly interesting hypothesis on which to focus, since no realistic model suggests it might be true. A question which is likely to become more pressing as time goes on, is not whether climate change is taking place, but what chance we have, given a 12-year data-set, of narrowing the range of uncertainty in climate model predictions. Precisely the same analysis as before can be applied to the global SST data with trend, annual cycle and SOI component removed, this time testing the hypothesis that the trend is less than 0.015°C per year, to see at what stage we may expect the data to begin to allow us to reject the lowest rates of warming currently predicted by climate models. The result simply reproduces figure 10 with the lines shifted 0.015°C to the right (I trust the reason, by now, is obvious). We need ~11 years of data to have a 50% (80%) chance of rejecting the hypothesis that the warming rate is less than or equal to 0.015°C per year if the true warming rate is 0.025°C (0.04°C) per year.

We can apply the same methodology to investigating the number of years of data required to place an upper limit on the rate of warming. Figure 11 shows that, to have a 80% chance of rejecting the hypothesis the warming rate is greater than or equal to 0.04°C per year, we would need ~11 years of data if the true warming rate was 0.015°C per year (point *a*).

This section aimed to demonstrate that a data-set of the length likely to be

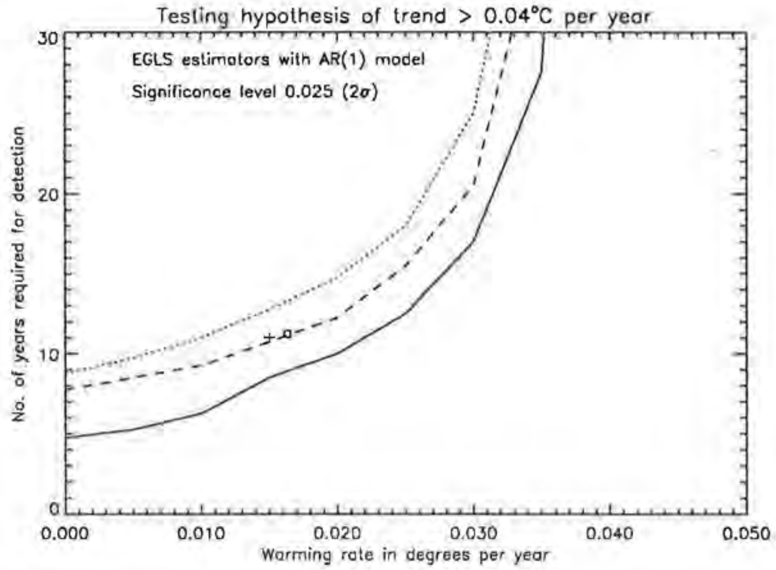

Figure 11: As previous figure, but considering the null hypothesis that the trend in global temperatures is *greater* than 0.04°C per year, the highest end of the range given by the IPCC.

provided by the ATSR series of instruments has a high chance of making a singificant contribution to narrowing the uncertainties regarding the rate and magnitude of global change. One final point must be made before we move on. We have assumed throughout that there is a negligible drift in the ATSR instruments. This point is discussed in the following section.

6 The impact of instrument drift on global change detection

If an instrument or system of instruments is to be used for the detection of a long term change, then the instrument(s) should clearly be as stable as possible over the time-frame in question. Even more importantly, we should be able to characterise and place objective and quantitative confidence limits on any residual instrument drift, or systematic error. It is this latter requirement which marks out the ATSR programme as particularly well suited, among the various observation systems available, to monitoring any global change in SST.

6.1 Quantifying drift in data from the ATSR and AVHRR instruments

As a self-calibrating radiometer, the short-term (month-to-month) time-scale drift in the ATSR is likely to be minimal. The instrument design and pre-flight calibration programme were specifically aimed at minimising the need to rely on *post hoc* corrections for non-linearities, which play an important role in the interpretation of AVHRR data. Such corrections are always based, at least implicitly, on a model of the measuring instrument. If they are not to introduce secular changes in measurement bias, it is important that the relationship between the true instrument characteristics and the characteristics assumed in this model does not change over time. Clearly, the more complex the model, the less confidence we have that this will be the case. The low level of pre-flight calibration, lower stability, and inherently greater non-linearity in the AVHRR instrument are the key reasons why it remains so difficult to use and rely on AVHRR data for climate research.

Reynolds *et al.* [33] demonstrate that the warming trend found in AVHRR-MCSST data by Strong [39] was almost certainly attributable to instrument biases and to the effects of the El Chicón volcanic eruption early in the series. Bates and Diaz, comparing AVHRR-MCSST data with surface observations through the 1980s, conclude that unsupported use of MCSST data for climate research was unjustified [6].

In the same study, Bates and Diaz also find that the surface observation data from ships-of-opportunity is inadequate even to resolve the annual cycle of SST south of about 40°S. They therefore advocate using blended MCSST and ship observations to obtain the required stability and global coverage. Use of blended data in this way is also problematic, because it introduces more possibilities for spurious drift than in data from a single instrument. The proportional weight given to the two data-sets may vary systematically over time, or the assumptions inherent in the blending procedure may break down. For example, the relationship between MCSST data and ship data may vary systematically both with location and with time: after the eruption of Mount Pinatubo, the stratospheric aerosol released was confined to the northern hemisphere for some months. During that time, AVHRR observations were seriously affected in the northern hemisphere, while remaining relatively unaffected south of the tropics. The standard blending procedure [30] effectively involves assuming that the MCSST data gives an accurate representation of SST gradients in areas which are not covered by surface observations, allowing these gaps to be filled in with satellite data "anchored" to the surface data where it is available. The presence of strong gradients in the concentration of a contaminant such as volcanic aerosol would render this assumption unjustified.

A more fundamental problem with the use of blended data from more than one source is that the two data-sets cannot then be used to cross-check each other. For example, there is a very good correlation between the blended analysis of Reynolds and Marsico [31] with the ship-only record of Bottomley *et al.* [8], but Folland *et al.* [13] note that this correlation is relatively meaningless, since the two data-sets contain much common data. The ATSR series of instruments should significantly improve this situation, since the dual-angle view appears to be able to deliver adequately accurate SST observations despite the presence of volcanic aerosol, with near-global coverage. Smith and Saunders [36] found the dual view to be crucial in compensating for the effects of aerosols released by Mount Pinatubo. In contrast, Reynolds [32] documents the substantial degradation of the (single view) AVHRR data through the period following the Pinatubo eruption.

As well as being able to check for any drift due to changes in the instruments' characteristics and overcome the effects of relatively short-term events like volcanic eruptions, we also need to be able to check for a drift in satellite-measured radiances due to a secular change in atmospheric thermal emission and absorption properties. Such a change might well be caused, for example, by a moistening of the lower troposphere, which is predicted to be one of the consequences of global warming. Such a moistening would itself be an important indicator of global change. An ability to distinguish between a secular change in SST, a drift due to instrument characteristics, and a drift due to secular changes in non-temperature climate variables, is clearly essential if an instrument is to be useful for global change research.

The dual-angle view will also be important here, since it allows the effect of the atmosphere to be characterised much more precisely than is possible with a simple nadir sounding instrument. The fact that no significant bias is found between skin SST as measured by the ATSR and as measured by a ship-mounted radiometer indicates that the present algorithm is successful in compensating for the effects of atmospheric water vapour, even with a relatively crude model of the atmospheric moisture field (see below). No such check can be made on AVHRR data, since the algorithms used to convert AVHRR satellite radiances to SST rely on an empirical regression between radiances and *in situ* buoy observations of bulk SST. No attempt is made to derive surface skin temperature from AVHRR data using physical principles. AVHRR retrievals deal with the effects of the skin-bulk temperature difference, atmospheric emission and absorption and changes in characteristics between instruments within in a single algorithm. This makes it impossible to distinguish between an increase in atmospheric absorption and a change in instrument characteristics using AVHRR data.

Work is in progress, however, on actually retrieving information on lower-tropospheric water vapour from ATSR data. It has already been shown (Barton *pers. comm.*) that the required information may, in principle, be contained in the ATSR data, although algorithms to extract it have yet to be developed and validated. Once this is achieved, it will be possible to perform a simultaneous retrieval of lower tropospheric water vapour and SST from the ATSR data. Better still, the water vapour field from an operational weather analysis like that of the ECMWF can be used as the first guess in the retrieval, making it, in effect, a one-dimensional variational data assimilation problem [12]. This cannot but improve the accuracy of the retrieved SST. Thus the current validation results on the ATSR should be seen as an upper limit on the instrument's accuracy, since they are based on a retrieval algorithm which, although physically based, assumes an atmospheric water vapour profile which varies only with latitude, not with longitude or time.

Any drift on longer time-scales can be monitored and, if necessary, corrected for by calibration against *in situ* measurements of skin SST made by ship-mounted radiometers. This is clearly a priority for the development of a reliable SST data-base for climate research purposes. A project is already in progress in RAL to develop a robust and accurate radiometer to be mounted on the masts of research vessels to provide "ground truth" observations of surface radiometric SST for ongoing ATSR validation work. A few such radiometers are in existence and have been extensively used in the initial validation exercise, but extending surface observations to give a wider geographical coverage and range of atmospheric conditions is clearly a priority.

A key characteristic of the ATSR project is that we are able to quantify and characterise any long-term drift in this way. The implications of such a validation exercise would be much less clearly defined for the AVHRR instrument. Since individual AVHRR instruments have their own retrieval coefficients based on an empirical regression on buoy data, then any monitoring for long-term drift in AVHRR data would, in effect, be monitoring for drift between the validation data-set and the buoy data. Such an exercise would be useful to check these data-sets, but it is not clear what it would mean for the AVHRR.

6.2 Quantifying drift in ship-of-opportunity data

Monitoring any secular change in systematic error in ship-of-opportunity observations is even more complicated, because we are not then dealing with a single measuring instrument. Trenberth and co-workers [42, 43] have investigated sources of error in *in situ* data from the COADS dataset, which is based on ship-of-opportunity observations. He identifies a number of sources of error, including

- Errors in individual observations, which includes the errors introduced by the different observation systems used. There is a significant bias between engine intake temperatures and bucket measurements, which also depends on the size of the engine intake and its depth below the surface. The concern for long-term climate monitoring is that a secular change in average ship design (which might be caused by an actual change, or simply by bringing a higher proportion of smaller vessels into the observation network in an attempt to improve coverage) could introduce a spurious drift in reported engine intake temperatures, just as the change in bucket design created the need for the well-known bucket-corrections in ref. [8].

- Incomplete sampling of the diurnal and seasonal cycles, and incomplete sampling of spatial gradients within a $2^\circ \times 2^\circ$ ($1^\circ \times 1^\circ$ in ref. [8]) grid square. Again, this could introduce spurious trends should trading routes and popular sailing times alter systematically during the period of monitoring.

Trenberth estimates the cumulative effect of these various sources of error to be an anticipated 1.2°C standard error in the observation of a monthly mean SST in a $2^\circ \times 2^\circ$ grid square, which is confirmed by the results of Harrison and Jones [18]. In contrast, the observed standard error between ATSR and buoy data in a daily observation on a $0.5° \times 0.5°$ grid is less than 0.5°C. An average of 240 such observations (assuming the day-time and night-time coverage achieved by ATSR to date) would contribute to a monthly mean observation on a $2° \times 2°$ grid. The standard error would not, of course, be reduced by a factor of $\sqrt{240}$, since many of the sources of noise will be autocorrelated, but the final result would clearly be much less than 1.2°C.

The key point to note here is not so much size of the errors in the ship-of-opportunity data, but the difficulty of characterising and quantifying these errors in order to check for the presence of a long-term systematic drift. Trenberth uses a variety of *post hoc* methods based on an analysis of the statistics of the ship reports. This is the best that can be done, given that we have very little control over those making the original observations. But it is clearly vastly inferior, in terms of the confidence we have in the derived confidence intervals, to a systematic analysis of error sources and validation against an independent data set, such as is possible with the ATSR. With the ATSR, we can predict the error variance on the basis of the instrument's design characteristics and uncertainties in the retrieval algorithm. If the observed error against a validation data-set is within the predicted error, we can conclude with some confidence that the instrument is working and measuring what we think it is measuring. With the ship data, we have to devise, often quite complicated, models of error sources, and generally the only way we have of checking whether these models are adequate is whether they account for the statistics of the data themselves. This can be quite a sophisticated art, and is clearly essential if we are to make use of ship data from past decades. But it leaves open the possibility of an unexpected source of error going undetected.

Finally, exclusive reliance on the ship-of-opportunity observations carries a strong risk of a significant change going undetected simply because it occurs in the southern hemisphere. Trenberth concludes

> With the marked exception of the eastern tropical Pacific ... there are insufficient numbers of *in situ* SST observations to reliably define SST or surface air temperature monthly mean anomalies over most of the ocean south of about 10°N. ... For seasonal means, SST anomalies cannot be reliably defined south of 20°S in the eastern Pacific, and south of ~35°S elsewhere except near New Zealand.

The implications of this problem are discussed in the following section.

7 Impact of spatial and seasonal observational coverage on fingerprint detection of climate change

The usual argument put forward in favour of supplementing surface observations with satellite data is the vastly improved spatial coverage provided by satellites. We are, however, unable to locate an exact comparison in which satellite and ship data have been treated in precisely the same way. Figures 12 to 15, showing work carried out with Mike Panter at RAL, attempt to redress this situation. Following the procedure of Bottomley *et al.* [8], updated with the documentation accompanying the UKMO ATLASC data-set, we begin by averaging all ATSR ASST observations which lie within a $1° \times 1° \times$ 5-day spatio-temporal location and which satisfy simple quality control criteria (lying within the range -2°Cto 37°C; within 6°C of climatology) to give a single "super-observation". Bottomley *et al.* express this as an anomaly about the climatology for that location and month in the year: with only two years of ATSR data, we cannot do this, but it does not affect our figures on spatial coverage. Trenberth remarks that a $1° \times 1° \times$ 5-day resolution is approximately the minimum allowable for SSTs not to be significantly affected by undersampling of spatial and temporal gradients.

Figures 12 and 13 show the number of such "super-observations" made by the ATSR instrument in a $5° \times 5°$ grid square in the months of January and July 1992 respectively. Most regions achieve between 100 and 200 super-observations, out of a maximum possible of 300 (25 $1° \times 1°$ squares in a $5° \times 5°$ square, and 6 pentads in a month, with both day-time and night-time data treated as independent observations.). Producing the equivalent figure for the ship-of-opportunity data-set would be very informative, but the necessary data is not available to the author of this report at this time.

Bottomley *et al.* calculate a $5° \times 5° \times$1-month average SST anomaly if at least 2 (out of a possible 150) such "super-observations" are defined in a given month. Figures 14 and 15 show the proportion of months covered by ship-of-opportunity observations for the months of January and July during the period 1980-1989 (from the MOHSST5 data-set [8], updated to version 5 in October 1992. Data provided by D. E. Parker and M. Jackson, UKMO). Coverage is clearly erratic in the Southern Hemisphere, particularly in the southern winter, even if we accept only 2 SST observations as defining a $5° \times 5° \times$ 1-month average.

Figures 16 and 17 show the equivalent coverage for the ATSR instrument. To compensate for the fact that we have treated day and night-time ASSTs independently, we require at least 3 ATSR super-observations in a $5° \times 5°$ grid square to define a monthly anomaly. Coverage is virtually complete, except in regions where we would expect large amounts of sea ice.

Because we are only two years into the ATSR mission, only sample months can

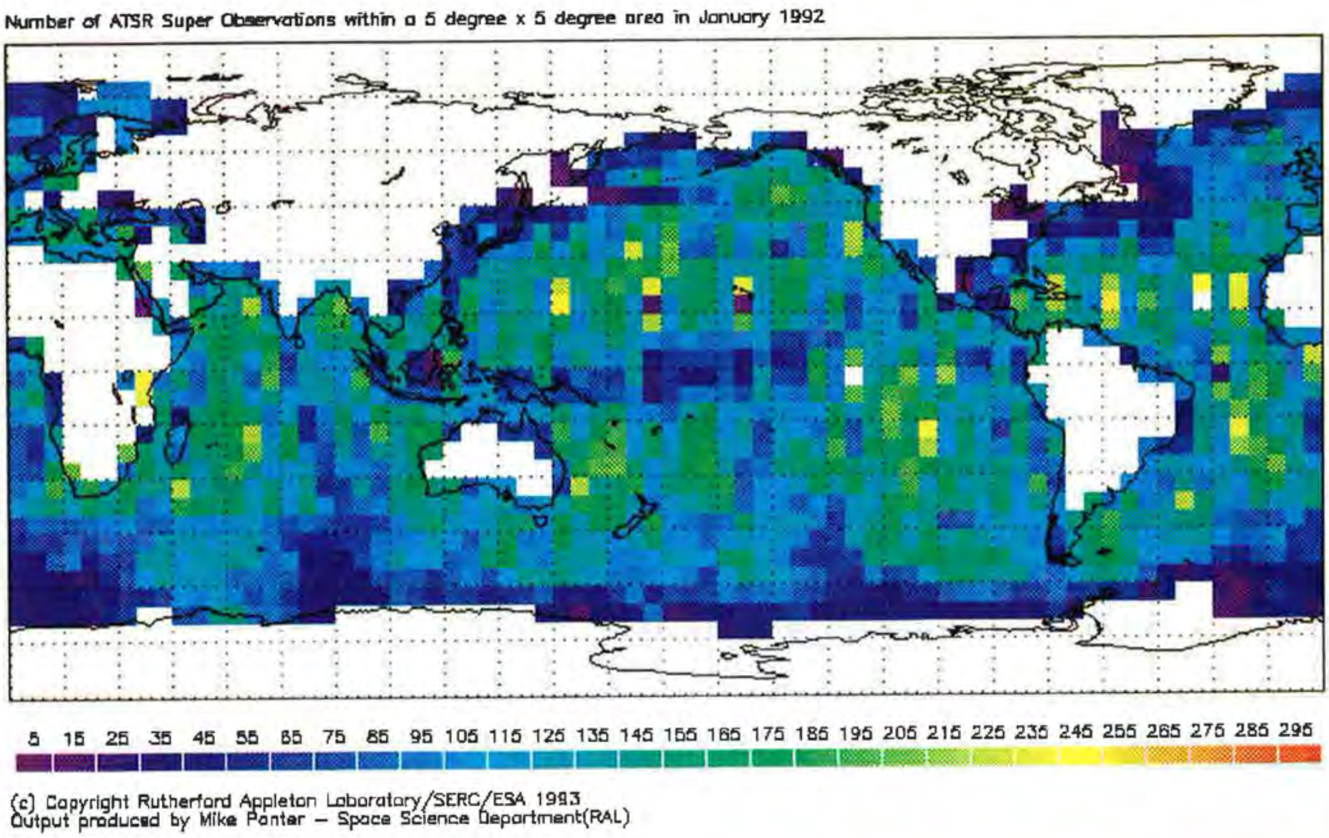

Figure 12: Number of "super-observations" (successfully observed $1^\circ \times 1^\circ \times$ 5-day spatio-temporal locations) made by the ATSR in a $5^\circ \times 5^\circ$ square as a function of the position of the square for January 1992. The maximum possible number is 300. White corresponds to no observations.

be provided for the satellite coverage, but there is no reason to believe these are unrepresentative for an instrument on a polar-orbiting platform. Likewise, there is no reason to believe that the ship-of-opportunity coverage is likely to improve significantly from that achieved during the 1980s, unless substantial resources are devoted to deploying research vessels or tethered buoys in the Southern Ocean, since this is largely determined by the pattern of trading routes.

The key point to note from these figures is that the satellite provides near-complete coverage of the region 70N–70S at this spatio-temporal resolution, while large proportions of the Southern Ocean and South Pacific are not consistently covered by the ship-of-opportunity observations. Before assessing the potential impact of this incomplete coverage on the detection of global change, we need to clarify exactly what it is we are trying to detect. The simplest possible manifestation of a global climatic change would be a uniform warming, i.e. (if we are focussing on SST) a change in SST which is independent of season and geographical location. The optimal way to detect such a change in a noise-contaminated data-set in which the noise amplitude is also independent of location and season would be to take a simple annual area-mean.

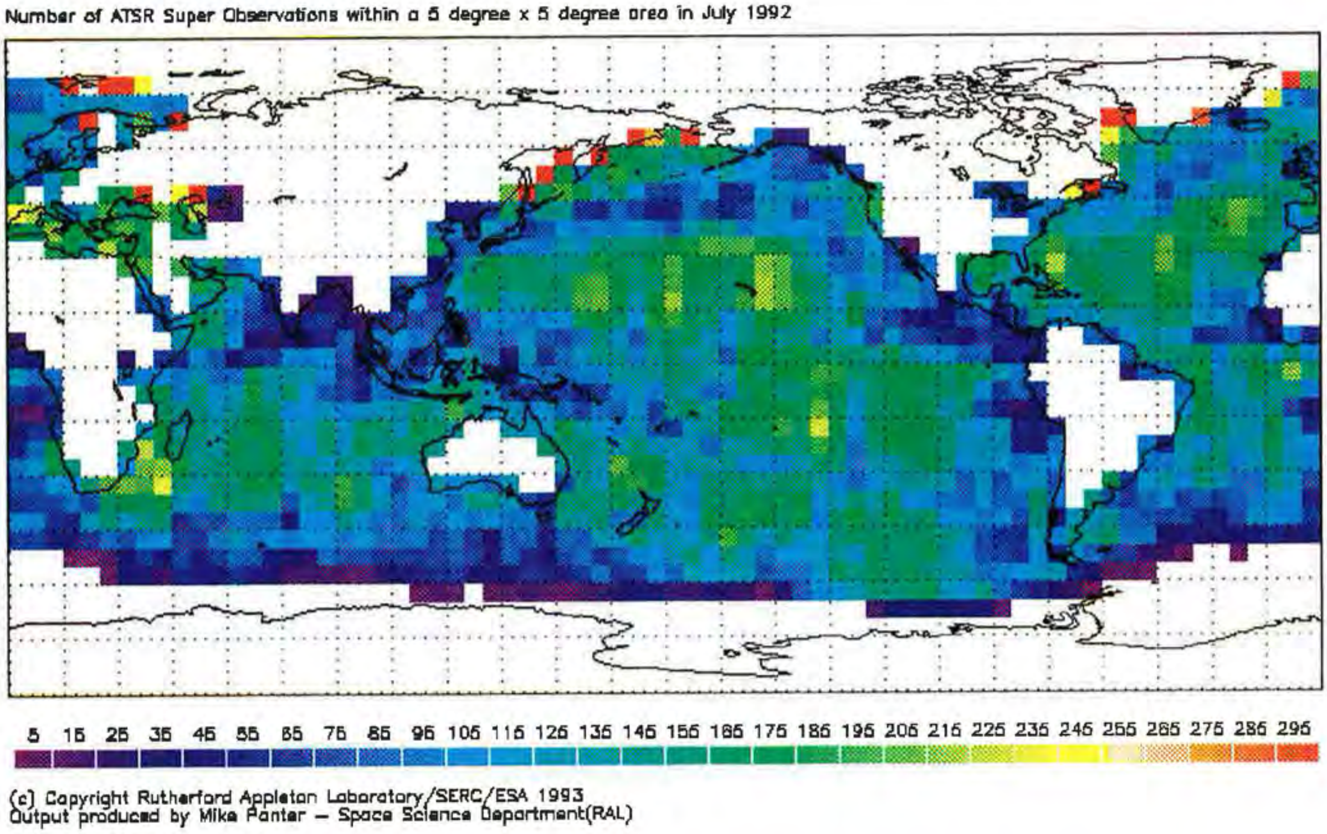

Figure 13: Number of "super-observations" (successfully observed $1^\circ \times 1^\circ \times$ 5-day spatio-temporal locations) made by the ATSR in a $5^\circ \times 5^\circ$ square as a function of the position of the square for July 1992. Note the greater extent of ice reducing coverage around Antarctica.

Taking into account the variation with latitude of the area of a square on a $5^\circ \times 5^\circ$ grid, this is given by

$$\overline{T}(t) = \frac{\sum_{x,y,m} T(x,y,m,t) w(x,y,m,t) \cos(y)}{\sum_{x,y,m} w(x,y,m,t) \cos(y)} \tag{2}$$

where $T(x,y,m,t)$ represents SST at longitude x and latitude y in month m of year t. $w(x,y,m,t)$ is a weighting factor which is equal to 1 if an SST observation was recorded for that particular location, month and year, and 0 otherwise. If the noise is correlated in space and/or time, or the noise amplitude is non-uniform, then the efficiency of $\overline{T}$ as an estimator of the "true" (noise-free) mean of the data may be improved by introducing more complex weighting functions on the RHS of equation (2). Since the purpose of this study is simply to investigate the impact of spatial coverage on detection-times rather than attempt any actual detection, we will not attempt a detailed review of such issues.

There are two problems with $\overline{T}(t)$ as a detection statistic: first, a global climatic change with important socio-economic implications, such as a warming of high

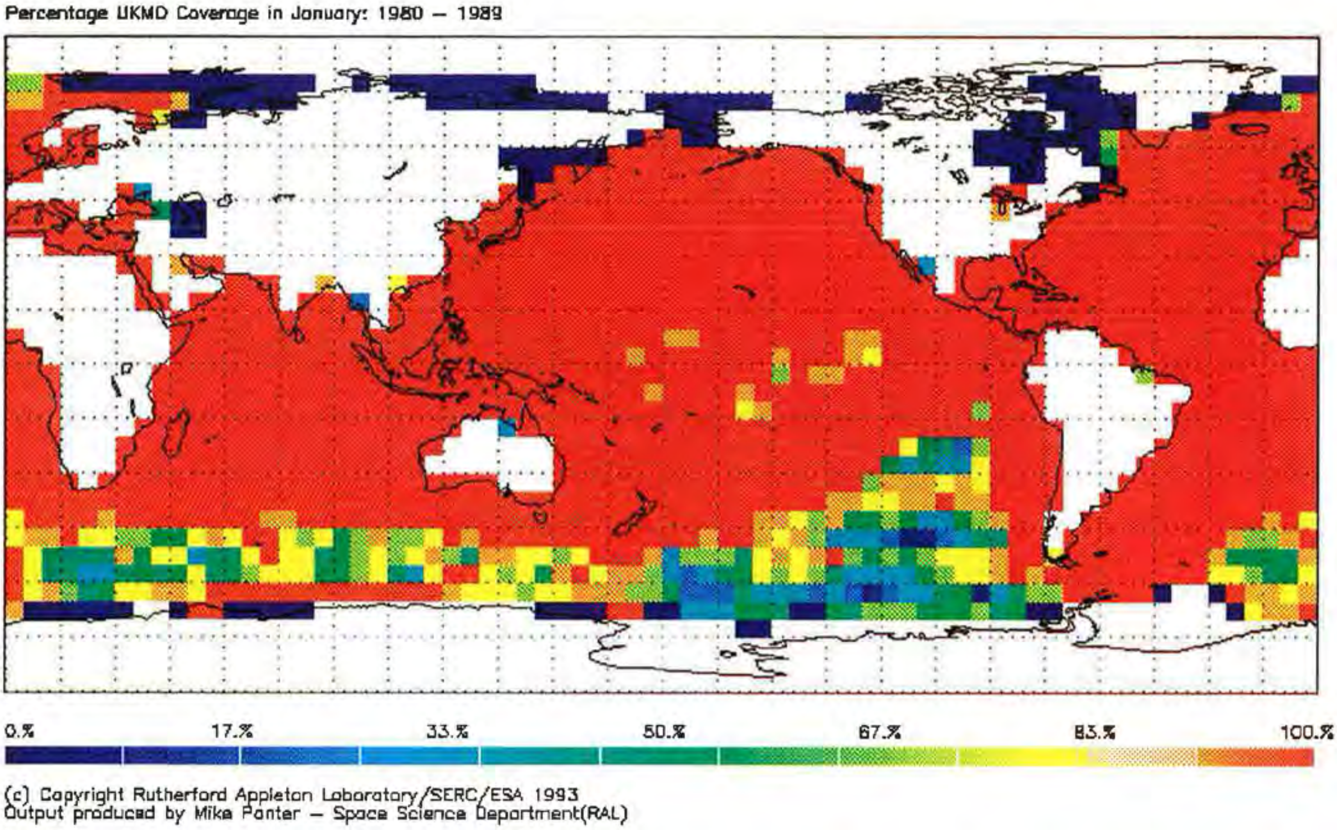

Figure 14: Proportion of Januarys during the 1980s in which at least 2 "super-observations" (out of a possible 150) appear in a $5^\circ \times 5^\circ$ grid square in the MOHSST5 ship-of-opportunity data-set

and mid-latitude regions accompanied by a corresponding cooling of the tropics, might have an insignificant impact on $\overline{T}(t)$. In other words, there might be a significant climate change without any accompanying global warming. This is an extreme scenario: current models consistently indicate that a change in $\overline{T}(t)$ will be one component of GHG-induced climate change. However, averaging temperatures over the globe and simply inspecting $\overline{T}(t)$ must involve discarding a great deal of information, some of which might allow us to detect climate change sooner than we would do if we relied solely on $\overline{T}(t)$.

Second, a systematic drift in measuring instrument characteristics might introduce a spurious trend into $\overline{T}(t)$, making it difficult to *attribute* an observed change in $\overline{T}(t)$ to the enhanced greenhouse effect. Both these problems may be overcome if we use prior knowledge, gained from physically-based climate models, of the characteristics of the signal which we are looking for, to develop a more specific detection statistic than $\overline{T}(t)$. This approach is known as statistical fingerprinting.

Statistical fingerprinting, in essence, involves looking for a pattern of change rather than simply a change in a single number like the global mean temperature. The

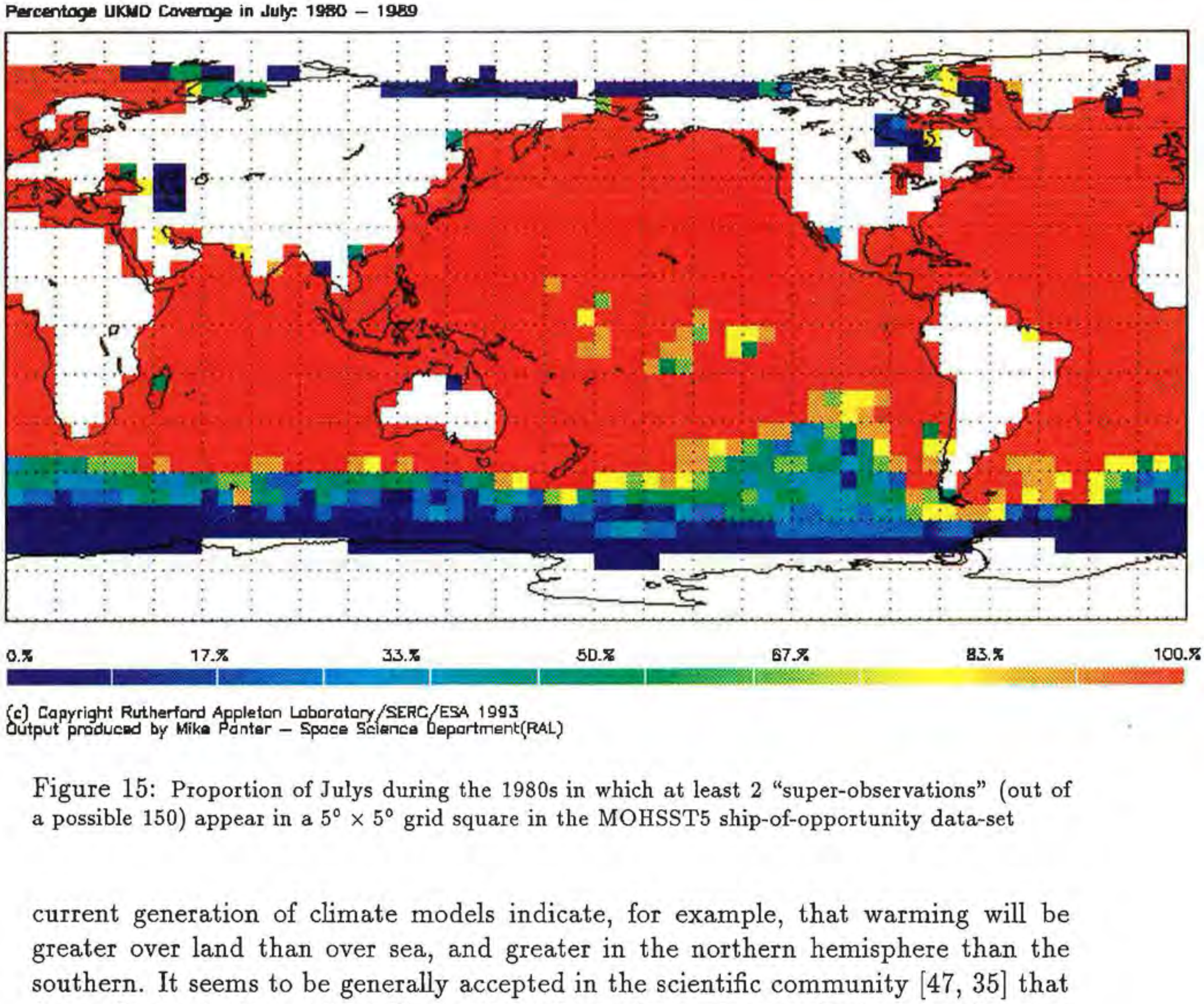

Figure 15: Proportion of Julys during the 1980s in which at least 2 "super-observations" (out of a possible 150) appear in a $5° \times 5°$ grid square in the MOHSST5 ship-of-opportunity data-set

current generation of climate models indicate, for example, that warming will be greater over land than over sea, and greater in the northern hemisphere than the southern. It seems to be generally accepted in the scientific community [47, 35] that for a change to be attributable to the enhanced greenhouse effect, not only must the change in $\overline{T}(t)$ be of the correct sign and magnitude, but the overall pattern of change must also be similar to the pattern predicted by climate models forced with increasing GHGs. If a global warming were to be observed which was much more pronounced over sea than over land, we might be inclined to suspect either that whatever instrument(s) were being used to measure SST were subject to a positive drift, or that whatever change was taking place was not that which was predicted by the climate models.

Insisting on the "correct" pattern of change cannot, of course, be an absolute stipulation. If the warming were eventually to become so strong that it exceeded all credible sources of instrument error, but the pattern of change remained inconsistent with that predicted by climate models, we would have to conclude that the models must have mis-represented something important, and simply got the pattern wrong.

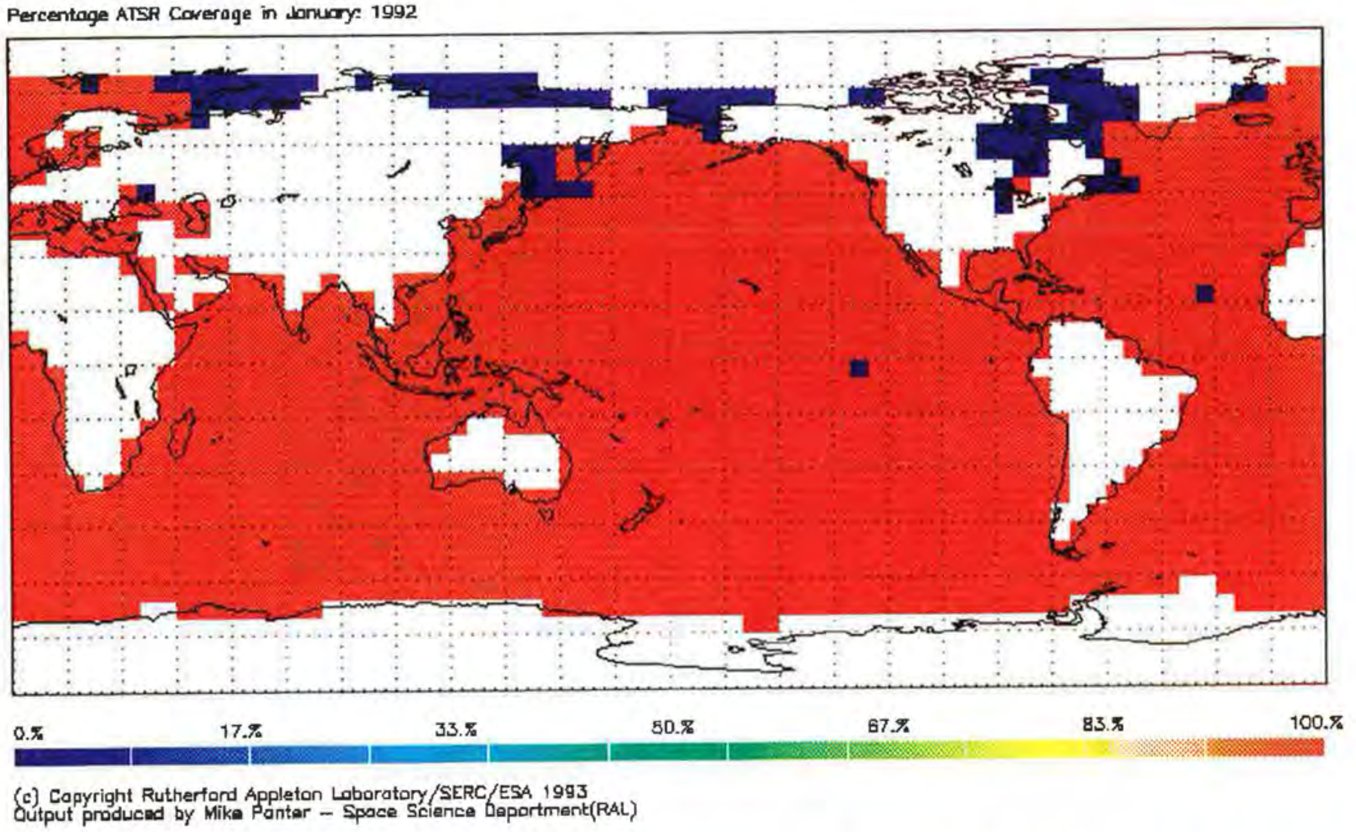

Figure 16: Coverage achieved by the ATSR in January 1992. Squares in which at least 3 "super-observations" (out of a possible 300) are shown as red, others blue.

At these early stages of detection, however, our confidence that we have detected a change which is genuinely attributable to the enhanced greenhouse effect would be considerably enhanced if the pattern of change were found to correspond reasonably closely to that predicted by the models.

The SST components of two such "fingerprints" from the UKMO Transient Response climate change simulation [25] are shown in figures 18 and 19, both corresponding to July. The difference between these two figures is that the first shows a simple fingerprint derived from the results of the increasing-CO_2 experiment, while the second shows the same fingerprint after an attempt has been made to correct for the drift in the control run the UKMO study. Both fingerprints have been "centered" (their annual mean components sets to zero), to ensure that an increasing projection onto such a pattern is not attributable to a change in the mean. The key point to note that significant structure appears south of 35°S, which would not be well observed by the ships-of-opportunity.

Our preliminary results may be summarised as follows: if all we are interested in is the detection of a significant change in $\overline{T}(t)$, then the additional coverage

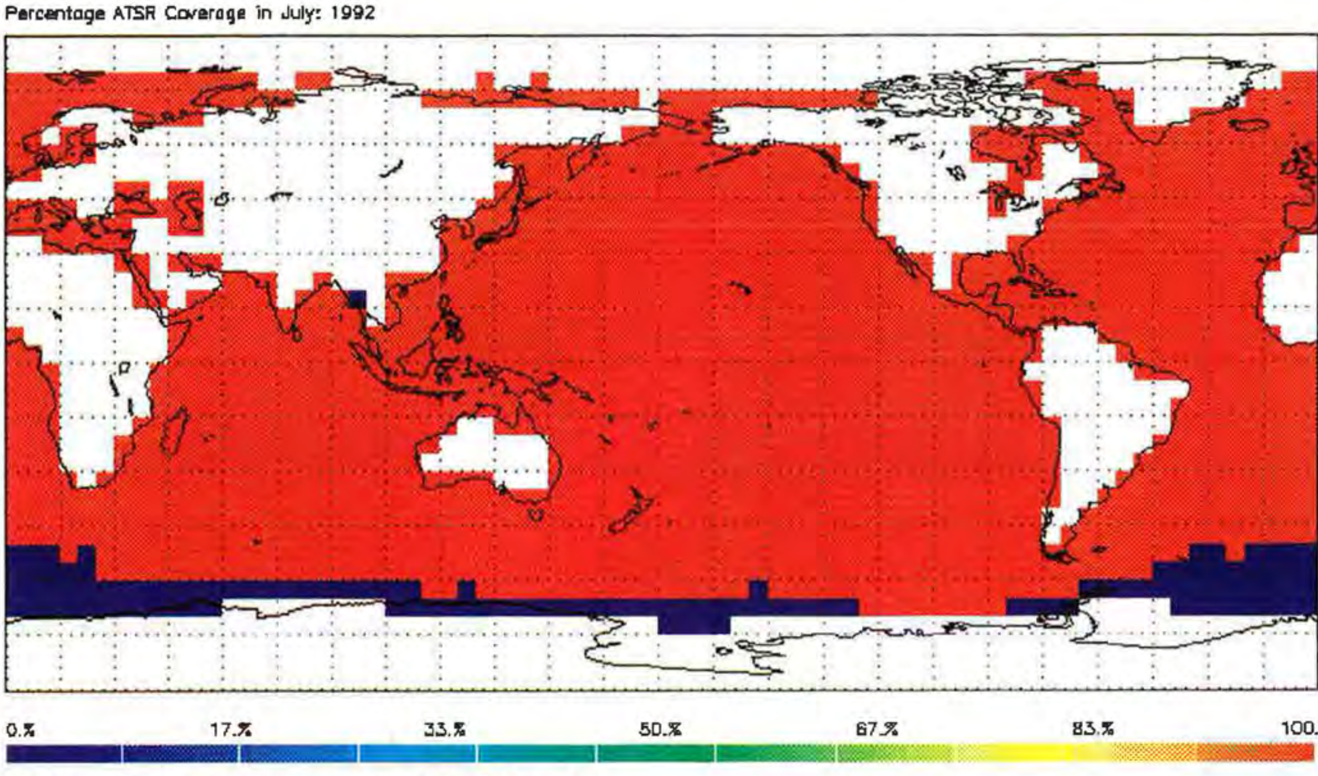

Figure 17: Coverage achieved by the ATSR in July 1992. Squares in which at least 3 "super-observations" (out of a possible 300) are shown as red, others blue.

offered by the satellites appears to make very little impact on the expected detection time. Simply detecting a change in $\overline{T}(t)$, however, is unlikely to satisfy the scientific community that GHG-induced climate change has begun: witness the fact that there has already been a significant warming trend in $\overline{T}(t)$ over the past century (although the significance of this trend does depend to some extent on the precise statistical model assumed), yet few climatologists are yet prepared to acknowledge that this trend is definitely attributable to the enhanced greenhouse effect. If, in addition to a significant trend in $\overline{T}(t)$, we also insist on detecting a significant trend in a centered fingerprint statistic (i.e. one which does not depend on $\overline{T}(t)$), then the additional spatial coverage afforded by satellites over ships-of-opportunity may have a much larger impact on detection time, although this result does depend on the details of the model-predicted pattern of change.

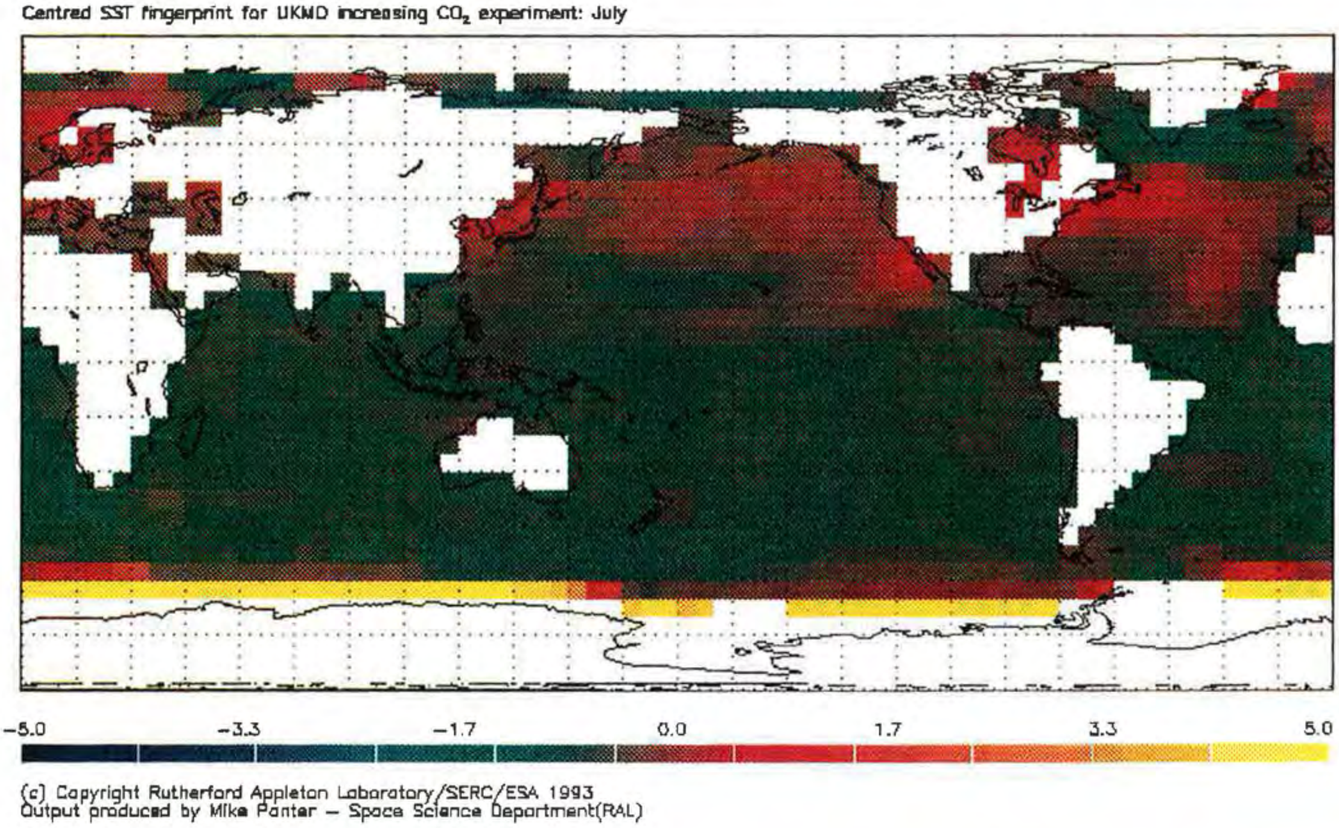

Figure 18: July component of centered "fingerprint" – pattern associated with warming trend – in the increasing-CO_2 run of the UKMO Transient Response climate change simulation. Note structure around the fringe of Antarctica, which would not be observed by ship-of-opportunity data.

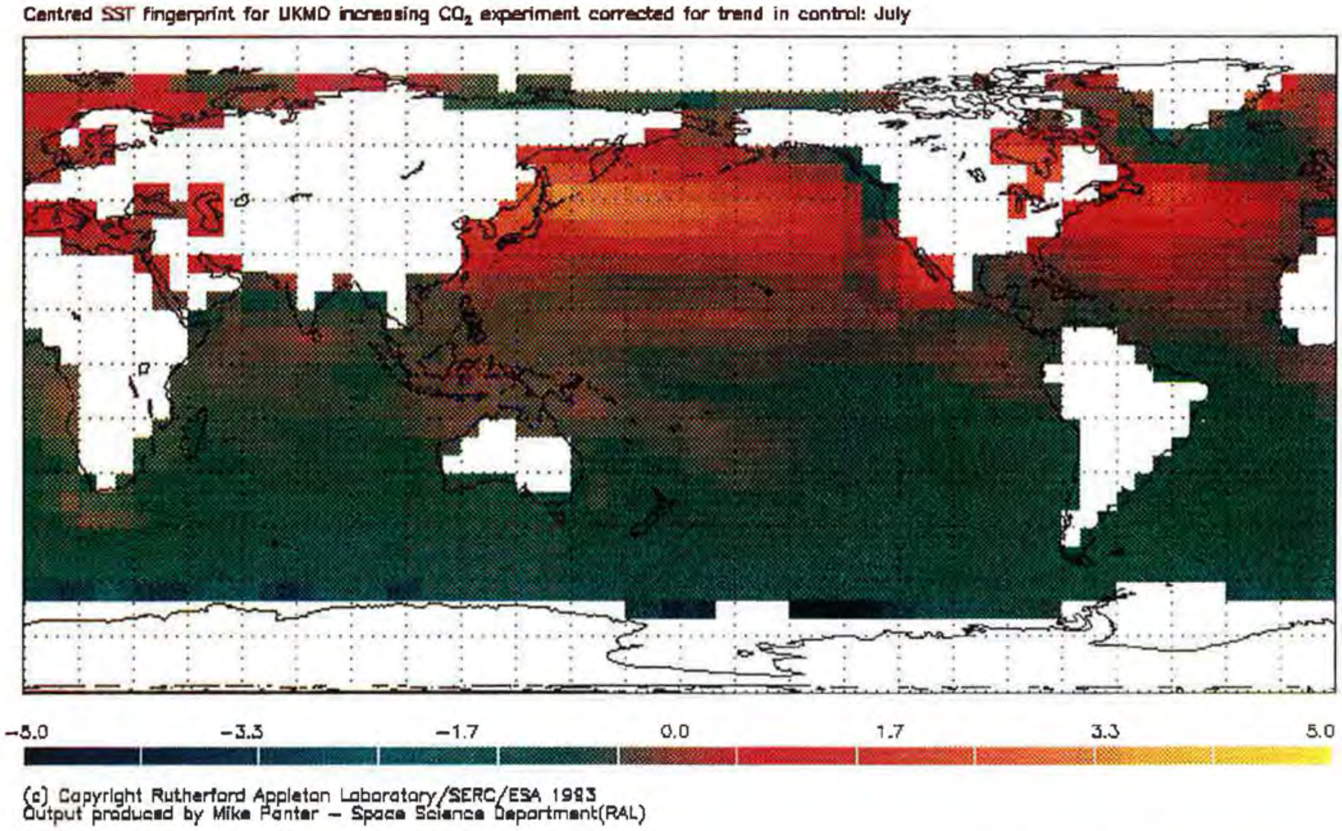

Figure 19: July component of centered "fingerprint" in the increasing-CO_2 run of the UKMO Transient Response climate change simulation after a correction has been made for the trend in the control. Note reversal of strong signal near Antarctica from previous figure, and introduction of a strong north-south dipole signature.

A Data-adaptive removal of locally-linear trends and cyclic components from a short time-series

This appendix describes a method for the removal of a trend and oscillatory components from a time-series. The key "data-adaptive" features of the procedure are that we do not assume that the trend is globally linear (i.e. we do not assume that it has the same gradient throughout the length of the series); and we allow for intermittent or gradual phase- and/or amplitude-modulation of the oscillatory component(s). The "paradigmatic" problem which this method is intended to address is the removal of trends and annual cycles (either natural or artificial) from geophysical data. It should be understood, however, that the method could also be applied to the removal of very-low-frequency stationary variability (treating it as a time-varying trend), as well as any cyclic component whose period is known *a priori*.

We emphasise that we do *not* attempt to address the problem of testing the statistical significance of trends or oscillations. We assume that the user wishes to reconstruct *any* trend or annual cycle, regardless of whether the data indicates it to be statistically significant. Tests are available, based on surrogate data [37, 41], designed to evaluate the significance of components extracted from a time-series in this way [4]. We strongly advocate their use if conclusions are to be drawn *directly* from the statistical properties either of the reconstructed trend and/or cycles or from the properties of the residual after the trend/cycles have been removed.

The method is derived from Singular Spectrum Analysis (SSA) [9, 14, 44] but avoids the problems of systematic bias in the estimated value of the trend near the end-points of a stationary series, as identified by ref. [1]. In a straightforward but (as far as we know) novel extension of SSA, we also allow for unbiased reconstruction of specific components of a time-series containing intermittent missing observations. We address the scalar problem here, although the extension to a vector series is conceptually straightforward.

The traditional method of removing a trend from data is a linear fit over the full series. This obviously gives rise to problems if the trend is non-linear, which might well be the case with historical data-sets if data-sampling and collection methods were particularly poor near the beginning of the series, but have since stabilised. Likewise, the standard method of removing an annual cycle is to assume a particular period to be "characteristic", calculate the mean annual cycle for that period to give a "seasonally-varying climatology", and subtract this climatology from the full series. This approach runs into problems if either the amplitude or the phase of the annual cycle component of the series in fact changes over time. This, again, might result from changing sampling patterns or observation procedures.

The method described here only assumes the trend to be linear, and the phase and amplitude of the annual cycle to be constant, over a user-prescribed time-scale: the "window width", M. This may be much shorter than the total length of the series. In standard SSA, problems of statistical estimation arise if the total time-series length is less than 3–4 times the window width [45]. Here, however, we are exploiting prior information in the form of our prior expectation of a locally linear trend (which SSA assumes in any case when applied to non-stationary series) and prior knowledge of the period of the annual cycle. This allows us to use any window width up to and including the full length of the time-series. In this latter limit, the method degenerates into a traditional, global fit to a straight line and regular sinusoid.

Consider an N-point time-series, $x_{t:\,t=1,N}$. We suspect this series to be contaminated with a trend which is approximately linear on time-scales less than or equal to M; and a cyclic component with (known) frequency f which is approximately sinusoidal but whose phase and amplitude may vary on time-scales greater than M. Generalisation of the method to deal with more than one cyclic component is straightforward.

Imagine sliding a window of width M down the length of the series. For each position of the window, we can estimate the local trend and the local amplitude and phase of the cyclic component by solving the following multiple regression problem, minimising $\sum_{i=1}^{M} \nu_i^2$ in

$$x_{i+j-1} = \sum_{k=1}^{4} e_{ik} p_{kj}^T + \nu_i \tag{A.1}$$

The index i indicates position within the window, $1 \leq i \leq M$, while j indicates the position of the window in relation to the original time-series. Specifically, j equals the index t of the first element contained in the window. If we stipulate that the overlap between the window and the series should always be complete, then there are $N - M + 1$ allowed window positions, hence[2] $1 \leq j \leq N - M + 1$. The $p_{jk:\,k=1,4}$ are coefficients to be estimated for this window position. The $e_{ik:\,i=1,M;\,k=1,4}$ make up a set of "basis vectors". They play an equivalent role to the EOFs of standard SSA, except that they are not assumed to be mutually orthogonal. For the reconstruction of a locally linear trend the required vectors are a zero-gradient-constant-mean straight line and a constant-gradient-zero-mean straight line:

$$\begin{aligned} e_{i1} &= r_1 \\ e_{i2} &= r_2 \left(i - \frac{M+1}{2} \right) \end{aligned} \tag{A.2}$$

where the r_k are normalisation factors. We choose r_k such that $\sum_{i=1}^{M} e_{ik}^2 = 1$ to maintain consistency with standard SSA (since we do not subsequently assume the basis to be orthonormal, the r_k can take any computationally convenient value). For a cyclic component of frequency f the required vectors are a pair of sinusoids in quadrature:

$$\begin{aligned} e_{i3} &= r_3 \cos \left[2\pi f \left(i - \frac{M+1}{2} \right) \right] \\ e_{i4} &= r_4 \sin \left[2\pi f \left(i - \frac{M+1}{2} \right) \right] \end{aligned} \tag{A.3}$$

[2] ...although C-programmers may prefer a different indexing convention.

Equations (A.2) are, of course, simply equations (A.3) in the limit $f \to 0$.

The vectors formed by the e_{ik} are not mutually orthogonal for all choices of M (i.e. $\sum_{i=1}^{M} e_{ki}^T e_{ik'} \neq 0$ for all $k \neq k'$), although they converge to orthogonality for large M. Thus we identify p_{jk} which minimise $\sum \nu_i^2$ in equation (A.1) using the standard approach of multiple linear regression, calculating a Hessian matrix and inverting it:

$$h_{kk'} = \sum_{i=1}^{M} e_{ki}^T e_{ik'} \tag{A.4}$$

$$p_{kj}^T = \sum_{k'=1}^{4} \sum_{i=1}^{M} h_{kk'}^{-1} e_{k'i}^T x_{i+j-1} \tag{A.5}$$

If some of the $x_{i+j-1:\, i=1,M}$ are missing observations, terms with the corresponding values of i are simply omitted from all summations over i, including the calculation of the elements of the Hessian, $h_{kk'}$. If a Jacobi-based method is used for the inversion [29], then the fact that some of the vectors made up by the e_{ik} may be mutually orthogonal can be exploited to minimise the computational cost of the procedure.

All we have done is "explain", in an ordinary-least-squares (OLS) sense, the behaviour of $x_{i+j-1:\, 1\leq i \leq M}$ in terms of a non-zero-mean straight line and a frequency f sinusoid with arbitrary amplitude and phase. In doing so, we have assumed implicitly that the rest of the behaviour of x can be modelled as an independent, gaussian-distributed "white noise" term. This will not, in general, be the case, but parameters estimated using OLS regression are, at least, unbiased for a wide variety of noise characteristics. OLS estimators are, however, not necessarily the most efficient available (in the sense of having the lowest possible variance) unless the white noise assumption is satisfied. Given that we will, subsequently, be averaging over large numbers of OLS estimators to obtain the reconstructed trend and cyclic component, this sub-optimal efficiency does not seem likely to have a significant effect in this application, but the extension of this work to accomodate auto-regressive (AR) or auto-regressive-moving-average (ARMA) processes, using generalised linear regression [11], would be an interesting subject for further study.

For this particular position of the window (value of j), we can reconstruct the trend and cyclic component of the series simply by writing down equation (A.1) omitting the noise term:

$$y_{ij} = \sum_{k=1}^{4} e_{ik} p_{kj}^T \tag{A.6}$$

The term y_{ij} represents an estimate of the trend plus cyclic component at time $t = i + j - 1$. We have up to M such estimates, corresponding to the positions of the window which overlap the point with time-index t. If all the data are of equal accuracy and there are no missing observations, all these estimates should be given equal weight, and an optimal single estimate of the trend and cyclic component at time t is given by their arithmetic mean:

$$r_t = \frac{1}{i_2 - i_1 + 1} \sum_{i=i_1}^{i_2} y_{i(t-i+1)} \tag{A.7}$$

where

$$\begin{aligned} i_1 &= \max(1, t - N + M) \\ i_2 &= \min(t, M) \end{aligned} \tag{A.8}$$

If there are missing observations, we should weight the terms $y_{i(t-i+1)}$ on the RHS of (A.7) by the number of observations used for the estimation of each term.

An obvious drawback of this technique, applied to the problem of removing an annual cycle, is that the cyclic component is assumed to be approximately sinusoidal. The simplest check on this assumption would simply be to calculate the annual cycle in the traditional manner and inspect it for any obvious departure from a sinusoid. If such a departure is observed, then it may be worth calculating a semi-annual cycle at the same time as the trend and annual cycle, by including two more basis vectors, with frequency (6 months)$^{-1}$, in the regression model. We clearly want to ensure that the total number of basis vectors used is always much less than M minus the maximum number of missing observations to occur within the window, or we risk "over-fitting", i.e. including either noise or signals not related to the trend or annual cycle into r_t. The choice of M will always be determined by a compromise between the need to maximise the amount of information used for the estimation of the local trend or annual cycle, and the need to minimise the length of time over which these are assumed to be constant.

A natural extension of this technique is to allow the e_{ik} to be determined by some other time-series, w_t, which is supposed to explain a proportion of the variability in x_t. If we have a model such as

$$x_t = b_1(t) w_t + \sum_{k=1}^{4} e_{ik} p^T_{k(t-i+1)} + \nu_t \tag{A.9}$$

and we assume that the coefficient b_1 varies only slowly over time, we simply introduce a fifth basis vector which, unlike the first 4, depends on j: $e_{i5} = w_{i+j-1;\, i=1,M;\, j=t-i+1}$. This gives a sliding-window based regression algorithm.

References

[1] M. R. Allen. *Interactions between the atmosphere and oceans on time-scales of weeks to years.* PhD thesis, University of Oxford, 1992. 202 pages.

[2] M. R. Allen. Biases in the estimation of trends and cyclic components near the end points of short time-series. *Geophys. Res. Lett*, 1994. in preparation.

[3] J. Murray, M. R. Allen, S. P. Lawrence, C. T. Mutlow and D. T. Llewellyn-Jones. Investigating intraseasonal oscillations in ATSR data. In *Proceedings of the Second ERS-1 Symposium*, 1993.

[4] M. R. Allen and L. A. Smith. Patterns of low-frequency climate variability. *Geophys. Res. Lett*, 1993. submitted.

[5] I. J. Barton, A. J. Prata, and D. T. Llewellyn-Jones. The Along-Track Scanning Radiometer - an analysis of coincident ship and satellite measurements. *Adv. Space Res.*, 13:569–574, 1993.

[6] J. J. Bates and H. F. Diaz. Evaluation of multichannel sea surface temperature product quality for climate monitoring: 1982-1988. *J. Geophys. Res.*, 96C11:20613–20622, 1991.

[7] L. O. Bengtsson. Climate system modeling prospects. In K. E. Trenberth, editor, *Climate System Modeling*, pages 705–729. Cambridge Univ. Press, 1992.

[8] M. Bottomley et al. *Global Ocean Surface Temperature Atlas.* H.M.S.O., London, 1990.

[9] D. S. Broomhead and G. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217–236, 1986.

[10] R. J. Charlson, J. Langner, H. Rodhe, C. B. Leovy, and S. G. Warren. Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic sulphate aerosols. *Tellus*, 43A-B:152–163, 1991.

[11] M. H. A. Davis and B. B. Vinter. *Stochastic modelling and control.* Chapman and Hall, 1985.

[12] J. R. Eyre, G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson. Assimilation of TOVS radiance information through one-dimensional variational analysis. Technical Report 186, ECMWF, 1992. 31 pages.

[13] C. K. Folland et al. Observed climate variability and change. In J. T. Houghton et al., editors, *Climate Change 1992, Supplement to the IPCC Scientific Assessment*, chapter C, pages 135–170. Cambridge Univ. Press, 1992.

[14] K. Fraedrich. Estimating the dimension of weather and climate attractors. *J. Atmos. Sci.*, 43:419–432, 1986.

[15] W. L. Gates, P. R. Rowntree, and Q. C. Zeng. Validation of climate models. In J. T. Houghton et al., editors, *Climate Change, The IPCC Scientific Assessment*, chapter 4, pages 93–130. Cambridge Univ. Press, 1990.

[16] M. Ghil and R. Vautard. Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, 350:324–327, 1991.

[17] J. Hansen et al. Global climate changes as forced by the Goddard Institute for Space Studies Three-Dimensional Model. *J. Geophys. Res.*, 93:9341–9364, 1988.

[18] D. L. Harrison and C. P. Jones. A user appraisal of ATSR near-real-time products. In *Proceedings of the First ERS-1 Symposium*, 1992.

[19] J. T. Houghton et al., editors. *Climate Change: the IPCC Scientific Assessment*. Cambridge Univ. Press, 1990.

[20] P. D. Jones. *Climate Monitor*, 17:80–89, 1989.

[21] J. E. Kutzbach. Modeling large climatic changes of the past. In K. E. Trenberth, editor, *Climate System Modeling*, pages 669–688. Cambridge Univ. Press, 1992.

[22] S. P. Lawrence, M. R. Allen, T. N. Stockdale, C. T. Mutlow, and D. T. Llewellyn-Jones. Investigating equatorial ocean dynamics in ATSR data and in an ocean model. In *Proceedings of the Second ERS-1 Symposium*, 1993.

[23] C. Lorius et al. The ice-core record: climate sensitivity and future greenhouse warming. *Nature*, 347:139–145, 1990.

[24] J. F. B. Mitchell and J. M. Gregory. Climatic consequences of emissions and a comparison of IS92a and SA90. In J. T. Houghton et al., editors, *Climate Change, The IPCC Scientific Assessment*, pages 173–175. Cambridge Univ. Press, 1990.

[25] J. Murphy. A prediction of the transient response of climate. Technical Report CRTN 32, Hadley Centre for Climate Prediction and Research, 1992.

[26] C. T. Mutlow et al. The Along Track Scanning Radiometer – validation results. *J. Geophys. Res.*, 1993. submitted.

[27] E. G. Njoku, T. P. Barnett, R. M. Laurs, and A. C. Vastang. Advances in satellite sea surface temperature measurement and oceanographic applications. *J. Geophys. Res.*, 90C6:11753–11586, 1985.

[28] A. J. Prata and I. J. Barton. A multi-channel, multi-angle method for the determination of infrared optical depth of semi-transparent high cloud from an orbiting satellite I: Formulation and simulation. *JAM*, October 1993. to appear.

[29] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: the Art of Scientific Computing.* Cambridge Univ. Press, 1989.

[30] R. Reynolds. A real-time global sea surface temperature analysis. *J. Climate*, 1:75–86, 1988.

[31] R. Reynolds and D. C. Marsico. An improved real-time global sea surface temperature analysis. *J. Climate*, 4, 1992.

[32] R. W. Reynolds. Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, 6:768–774, 1993.

[33] R. W. Reynolds, C. K. Folland, and D. E. Parker. Biases in satellite-derived sea surface temperature data. *Nature*, 341:728–731, 1989.

[34] J. E. Robertson and A. J. Watson. Thermal skin effect of the surface ocean and its implications for CO_2 uptake. *Nature*, 358:738–740, 1992.

[35] B. D. Santer, T. M. L. Wigley, and P. D. Jones. Correlation methods in fingerprint detection studies. *Climate Dynamics*, 8:265–276, 1993.

[36] A. H. Smith and R. W. Saunders. Validation of ATSR-1 using aircraft radiometer measurements over the South Atlantic. In *Proceedings of the First ERS-1 Symposium*, 1992.

[37] L.A. Smith. Identification and prediction of low-dimensional dynamics. *Physica D*, 58:50–76, 1992.

[38] R. W. Spencer and J. R. Christy. Precise monitoring of global temperature trends from satellites. *Science*, 247:1558–1562, 1990.

[39] A. E. Strong. Greater global warming revealed by satellite-derived sea surface temperature trends. *Nature*, 338:642–645, 1989.

[40] P. P. Tans, I. Y. Fung, and T. Takahashi. Observational constraints on the global atmospheric carbon dioxide budget. *Science*, 247:1431–1438, 1990.

[41] J. Theiler, S. Eubank, A. Longtin, B. Galdrikan, and J. D. Farmer. Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58, 1992. 77.

[42] K. E. Trenberth. Signal and noise in the surface temperature record. In *Preprints, 12th Conference on Probability and Statistics in the Atmospheric Sciences*, 1992.

[43] K. E. Trenberth, J. R. Christy, and J. W. Hurrell. Monitoring global monthly mean surface temperatures. *J. Climate*, 5, 1992.

[44] R. Vautard and M. Ghil. Singular Spectrum Analysis in nonlinear dynamics with applications to paleoclimatic time series. *Physica D*, 35:395–424, 1989.

[45] R. Vautard, P. Yiou, and M. Ghil. Singular Spectrum Analysis: a toolkit for short, noisy and chaotic series. *Physica D*, 58:95–126, 1992.

[46] G. A. Wick and W. J. Emery. A comprehensive comparison between satellite-derived skin and multi-channel sea surface temperature. *J. Geophys. Res.*, 97C4:5569–5595, 1992.

[47] T. M. L. Wigley and T. P. Barnett. Detection of the greenhouse effect in the observations. In J. T. Houghton et al., editors, *Climate Change, The IPCC Scientific Assessment*, chapter 8, pages 245–255. Cambridge Univ. Press, 1990.