

A COMPARATIVE ANALYSIS OF THREE LONG-TERM NDVI DATASETS DERIVED FROM AVHRR SATELLITE DATA

Keith R. McCloy¹, Sietse Los², Wolfgang Lucht³ and Søren Højsgaard¹

1. Danish Institute for Agricultural Sciences, Tjele, Denmark;
[keith.mccloy / soeren.hojsgaard}@agrsci.dk](mailto:{keith.mccloy / soeren.hojsgaard}@agrsci.dk)
2. University of Wales, Department of Geography, Swansea, United Kingdom;
s.o.los@swansea.ac.uk
3. Potsdam Institute for Climate Impact Research, Potsdam, Germany;
wolfgang.lucht@pik-potsdam.de

ABSTRACT

The AVHRR data provides the only long-term synoptic coverage that we have of the globe for use in global change and dynamics studies. At present three temporal datasets (Pathfinder, GIMMS and Fasir) have been developed from the AVHRR data. A major use of such datasets is in the analysis of trends that may be occurring in vegetation. The goal of the work reported in this paper was to conduct trend analysis using the data sets and to compare the results for similarities and differences. The work was conducted in two phases. The first phase derived trends in *NDVI* over the period of the datasets and compared these trends and their differences. The second comparison was to analyse the temporal correlation at each pixel between the datasets so as to better understand the discrepancies between the datasets.

All three of the datasets showed an average global trend of increasing *NDVI* over the period 1981 – 2000, with this trend varying between 1.8 – 4.5% depending on the dataset used. The correlation analysis showed that the Fasir and GIMMS datasets are more highly correlated than either are with the Pathfinder dataset. There are significant regional variations from this average global trend, with some of these variations being recorded in a consistent way between the different datasets, and some not. Some of these variations form strong linear anomalies, with the Norwegian and Sahelian linear anomalies being the most obvious. The regional differences between the datasets are of such a magnitude that it is considered that these datasets cannot be used for regional mapping or monitoring tasks that involve estimation unless the user is aware that quite different results are likely to be derived from the different datasets.

Keywords: Temporal image analysis, AVHRR, Pathfinder, Fasir, GIMMS, time series analysis, *NDVI*

BACKGROUND

The AVHRR series of satellites currently provide the only long term, daily coverage of the globe with data that are useful for the study of vegetation status. The AVHRR sensors started to acquire imagery useful for land applications in 1981. This acquisition continues to this day. The value of this relatively long record has been well recognised. It was used to study the dynamics of vegetation in Africa and the continental USA (T1,2). It has been used as evidence of earlier green-up of vegetation in the northern high latitudes (35°N to 75°N) (3,4). Other work has shown that an increasing trend in rainfall in the Sahel is matched by an increasing trend in the Normalised Difference Vegetation Index (*NDVI*) (5,6), given in Eq. (1) below. It has been used as the basis of continental and global landcover mapping (7,8) and others have used it to compare model estimates of green vegetation with the actual green vegetation status (9).

$$NDVI = \frac{\lambda_{NIR} - \lambda_{red}}{\lambda_{NIR} + \lambda_{red}} \quad (1)$$

In Eq. (1), λ_{NIR} is the response, radiance or reflectance in the near infrared and λ_{red} is the equivalent in the red. Whilst *NDVI* can be derived from either the response, radiance or reflectance values, the derived values will be different, simply because different transformation relationships exist between these attributes in each waveband.

The scientific value of this relatively long-term dataset has also been recognised by NASA, through their decision to make datasets derived from the data available to the scientific community through the Distributed Active Archive Centre (DAAC). The Pathfinder datasets of *NDVI* products at various spatial and temporal resolutions have been made available in this way for the period 1981 until 2000, when MODIS-based products superseded it.

The AVHRR data were never designed for many of the purposes to which they are being put, not the least for the study of long-term trends in global vegetation. As a consequence, many of the components necessary to create a calibrated, internally and externally consistent temporal data source were never established for the AVHRR series of satellites. This has given rise to a number of sources of lack of consistency in the data, including:

- the lack of in-flight calibration so as to deal with sensor deterioration,
- the drift of the time of acquisition to later in the day,
- the difficulty of conducting rigorous, after the event, atmospheric calibration for water vapour and aerosols.
- the wide swathe width creates difficulties with variations in radiance as a function of both path length and atmospheric conditions.

Nonetheless, the value of this relatively long-term record has meant that much effort has been and continues to be expended in attempting to derive products from the image data. These efforts have given rise to the development of three different approaches to the calibration and correction of the data in an attempt to produce internally and externally consistent datasets. The three datasets that have been developed are:

NASA Pathfinder (DAAC) dataset (10)

The source count data are corrected to radiance at the satellite using the pre-flight calibration constants modified by an approximation for degradation over time using the method of (11) that is based on the use of part of the Libyan desert as a standard surface. The radiance data are then converted to reflectance at the surface using the method of (12), by correcting for ozone absorption and Rayleigh scattering due to the atmosphere. The corrected data are then rectified to 8 km resolution and to 10-day composites by using the maximum *NDVI* value in each cell in the 10-day interval. The degree resolution, monthly data are then averaged again from the 8 km, 10 day data. The data have not been corrected for atmospheric moisture or dust contamination, both of which will reduce *NDVI* values in the datasets.

The International Satellite Land-Surface Climatology Project (ISLSCP) Fasir dataset

ISLSCP have supported the production of the Fourier-Adjustment, Solar zenith angle corrected, Interpolated and Reconstructed (Fasir) *NDVI* dataset (13,14,15,16,17). The approach with both the Version 2 Fasir and the GIMMS datasets is to derive reflectance at the top of the atmosphere, rather than at the surface. The Fasir dataset achieves this by filtering the data to remove pixels at extreme angles and areas affected by clouds. Data are then smoothed across each twelve monthly interval by fitting the data to a curve and adopting the fitted values for the twelve months of data. The data are then normalised for Solar Zenith angle, effectively providing data with nadir illumination and viewing directions. Daily *NDVI* is then calculated, the maximum value at each cell is then selected over the monthly period of acquisition and this value is corrected for sensor degradation in a similar manner to the DAAC dataset. The data are then averaged over the degree resolution elements. The data have not been corrected for atmospheric effects.

The Global Inventory Monitoring and Modelling System dataset (GIMMS) (18)

The GIMMS dataset was developed by NASA laboratories in Maryland using a similar philosophy to that of the Fasir 2 dataset, although the Fasir dataset is now supported by the University of Wales. The data are filtered for cloud contamination and the outer pixels are removed. Preflight calibration is conducted on the dataset to convert the data to radiance using the method described in (19), which is then converted to *NDVI*. The *NDVI* images are rectified and the maximum *NDVI* values selected in 10-day intervals to make a 10-day composite *NDVI* image. An approximate sensor correction is applied using the method described in (20). The data are neither corrected for atmospheric effects, nor normalised for variations in solar zenith angle.

An important question that arises concerns the differences that may arise in the results achieved when using one of these datasets rather than another of them in a scientific analysis. An objective evaluation based on field data is impossible in this situation, because of the lack of adequate historical data. Nor was it considered reasonable to conduct the analyses by comparing each dataset with other environmental variables, since the complexity of vegetation response will mask the comparison of the datasets. Vegetation responds to a release of constraints to growth; and since the constraints and their degree of constraint vary from location to location, such analysis will be complicated by the need to separate the vegetative component of the change from the component due to the different image sets used in the analysis.

The objective of this paper is to report on work done to compare the results achieved in using the different datasets in estimating trends in *NDVI* over the period 1981 – 2000. The second comparison is to investigate the temporal correlation between the datasets.

Caveat on this analysis

All of the work conducted in this analysis used Pathfinder data supplied from the NASA DAAC site in 1999. Dr C J Tucker provided the GIMMS dataset in February 2003. It was a test version that had not been generally released at that date. Dr Sietse Los supplied the Fasir 2 data in January 2003 and the Fasir 4 dataset in late 2003. Both the GIMMS and the Fasir datasets are undergoing continued refinement and so the datasets used in this analysis are not the latest datasets that are available. All of the data used in this study are the degree resolution, monthly datasets.

ANALYSIS OF THE LONG TERM TRENDS FOUND IN THE DATASETS

The data for each pixel were converted to annual average values, by taking the average of each twelve monthly observations in each year. These values were then used to derive both linear and quadratic curve fits by means of least squares. The linear curve fit did not produce statistically significant results at the level used in this paper, and as a consequence it is not further considered in this paper. Inspection of the Pathfinder data showed that it contained irregular errors at high latitudes in winter. As a consequence this part of the dataset was removed from the analysis, leaving a zone from 45°N to 23°S latitude in the analysis. There was also a gap in the Pathfinder dataset in late 1994; the July 1994 to June 1995 period was also removed from the analysis, and the gap treated as being a gap in the data in the analysis.

Statistically the analysis involved fitting the polynomial

$$y_t = \beta_1 + \beta_2 t + \beta_3 t^2$$

by least squares using the set of annual observation equations

$$Y = X\beta$$

in which $\beta = (\beta_1, \beta_2, \beta_3)^T$ is the vector of unknowns and T indicates transpose. The estimate of β is given by:

$$b = (X^T X)^{-1} X^T Y$$

To test the hypothesis, $\beta_2 = \beta_3 = 0$, a Wald test (21) was made. To account for correlation among observations at the same pixel, an autoregressive correlation structure was imposed on the covari-

ance array, Σ , such that each element of the array is the product of the variance, σ^2 and the correlation, ρ to a power that is a function of its position in the array. That is:

$$\Sigma_{t,t-k} = \sigma^2 \rho^k,$$

in which $|\rho| < 1.0$ so that the correlation decreases as k increases. Here $\sigma^2 = \text{Var}(y)$ while ρ is the correlation between successive years. The variance of b was estimated by:

$$V = (X^T X)^{-1} X^T \Sigma X (X^T X)^{-1}$$

To test the hypothesis, the χ^2 Wald statistic was calculated:

$$z^2 = (b_2, b_3) V_{2,3}^{-1} (b_2, b_3)^T,$$

in which $V_{2,3}$ is the sub-matrix of V that remains after removing the first row and column of V .

Table 1: Global trends in NDVI as derived from the four different calibrated datasets derived from NOAA AVHRR satellite image data for the period 1982 – 1999.

Dataset	All trends across all land surfaces			Trends significant at the 5% level, across all land surfaces			Trends significant at the 5% level across areas of significant trend.		
	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
Pathfinder	+3.27%	+17.3%	-13.5%	+1.81%	+17.3%	-13.5%	+5.81%	+17.3%	-13.5%
Fasir, V2	+3.94%	+86.9%	-28.5%	+2.47%	+86.9%	-28.5%	+8.50%	+86.9%	-28.5%
Fasir, V4	+8.24%	+167.7%	-56.2%	+5.06%	+148.5%	-56.2%	+11.46%	+148.5%	-56.2%
GIMMS	+5.81%	+108.1%	-33.0%	+4.51%	+108.1%	-33.0%	+9.49%	+108.1%	-33.0%

Software was developed to derive the regression, conduct the Wald test and return the z^2 value. The software also computed the trend as the ratio of the linear and quadratic terms over the period of the analysis, divided by the constant term. The time coordinates were set to zero at the start of the series, so that the constant term represented the estimated values from the regression at the start of the image series and the other terms will indicate the change over the period of analysis.

The trends derived from the quadratic fits are displayed in Figures 1-4. In each figure (a) represents all trends, whilst (b) represents the trends that are significant at the 5% level. The statistics derived from these images are given in Table 1. In these figures dark blue represents areas of water, snow and ice that were not analysed including Antarctica and most of Greenland, grey areas where the regression parameters were not significant at the 5% level, and the colours indicate the magnitude of the trend in accordance with the legend. In some instances there will be very good physical explanations as to why areas do not meet the 5% confidence level criterion. This can occur for areas that do exhibit significant deviations from the curve fit, due to the activities of man, or due to natural events. Thus in central Australia and eastern Queensland, the significant decreasing trends in *NDVI* are most likely due to the physical removal of woody vegetation, for quite different reasons in each area. Thus failure to meet the 5% criterion does not automatically mean that the data contains high levels of unexplained variation, or noise in the strict sense. In these areas there may exist spatial correlation between the pixels, however such analyses has not been conducted in this study. All of the datasets provided a positive trend over the period of analysis as can be seen from the summary in Table 1, but with noticeable regional fluctuations about this trend, as can be seen in Figures 1-4. There are extensive regional areas in the northern hemisphere with increases of greater than +10%, and regional areas of increase and decrease in the southern hemisphere. There are isolated areas of decrease in the northern hemisphere, where the most obvious is the decrease of about -4% (Fasir 2) in the eastern Canadian shrubland /needleleaf forest zone (22). Isolated areas of large negative trends in South America and areas of large positive trends in western Australia dominate the southern hemisphere.

The data in Column 2 in Table 1 show that the Fasir 4 dataset gave the largest mean trends, followed by the GIMMS dataset. A comparison of these two images (Figures 3(a) and 4(a)) shows that there are similarities but also significant differences. The extensive area of large increasing trend in Scandinavia and northern Russia that is so obvious in the GIMMS-based result does not exist in that derived from the Fasir 4 data. On the other hand, the Fasir 4 exhibited significantly higher trends in north central Siberia, larger areas of increasing trends in the Sahel and larger regional areas of negative trends in South America.

The trends derived from the Fasir 2 data are similar in location to that from the Fasir 4 dataset, but usually of lower values, so that the areas of large increasing trend are smaller. The notable exception to this is the central Siberian plateau, which does not exhibit large areas of large increasing trends in the Fasir 2 dataset. All three of the images exhibited larger areas of large increasing trends than did the Pathfinder data.

The statistically significant trends (column 5 in Table 1) show the same sequence as for the trends discussed above. The similarities between the two versions of the Fasir data are obvious, and the large area of high trends in Scandinavia, northern Russia and Siberia stands out in the GIMMS data in contrast to the Fasir datasets. Since global warming is expected by some to create a strong response in the temperature and radiance-constrained boreal zone, use of the GIMMS or the Fasir datasets might result in significantly different conclusions about the magnitude of this response. As a consequence, there is a need to try and better understand the magnitude and location of the discrepancies that exist between the datasets.

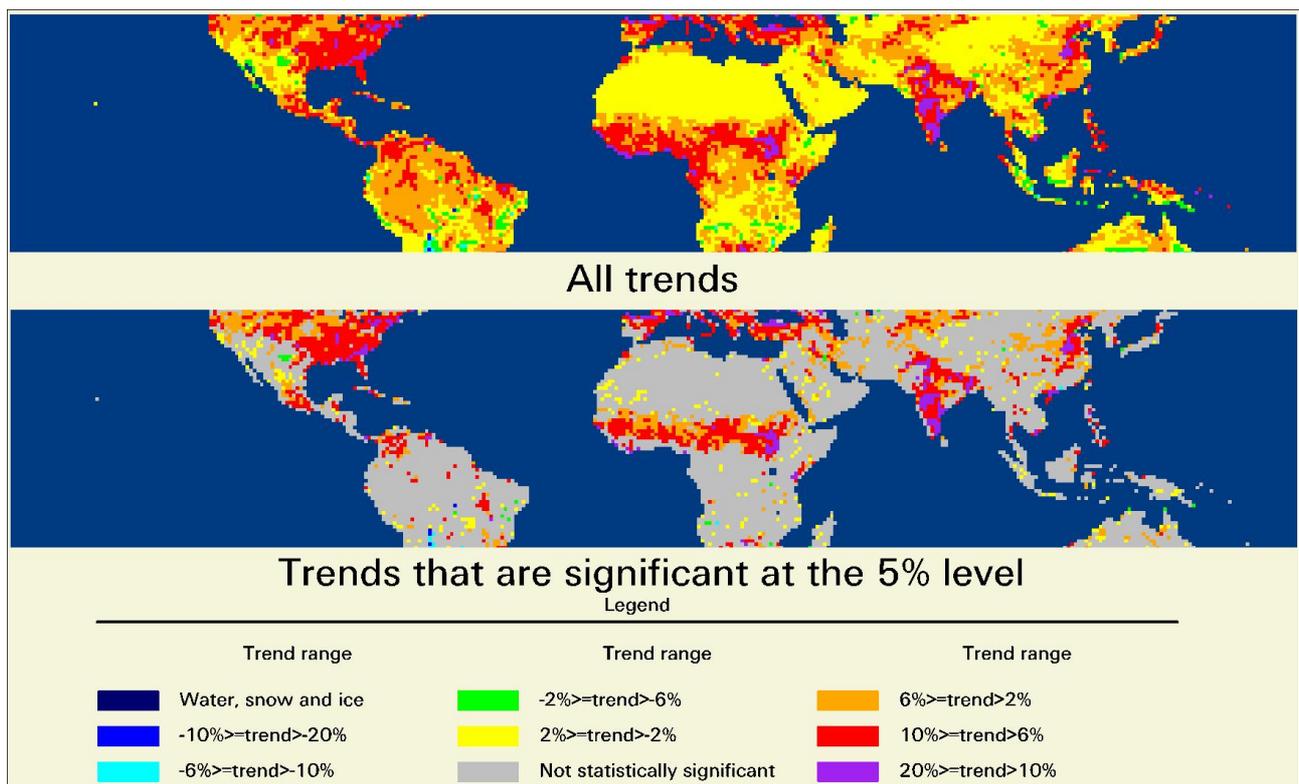


Figure 1: Percentage trends in NDVI as estimated from a quadratic least squares regression for the Pathfinder dataset, for the period July 1981 – June 2000, excluding July 1994 – June 1995, between 45°N and 23°S (18 years). In the figure, (a) represents all trends, whilst (b) are only those trends that are significant at the 5% level as estimated using the Wald Test.

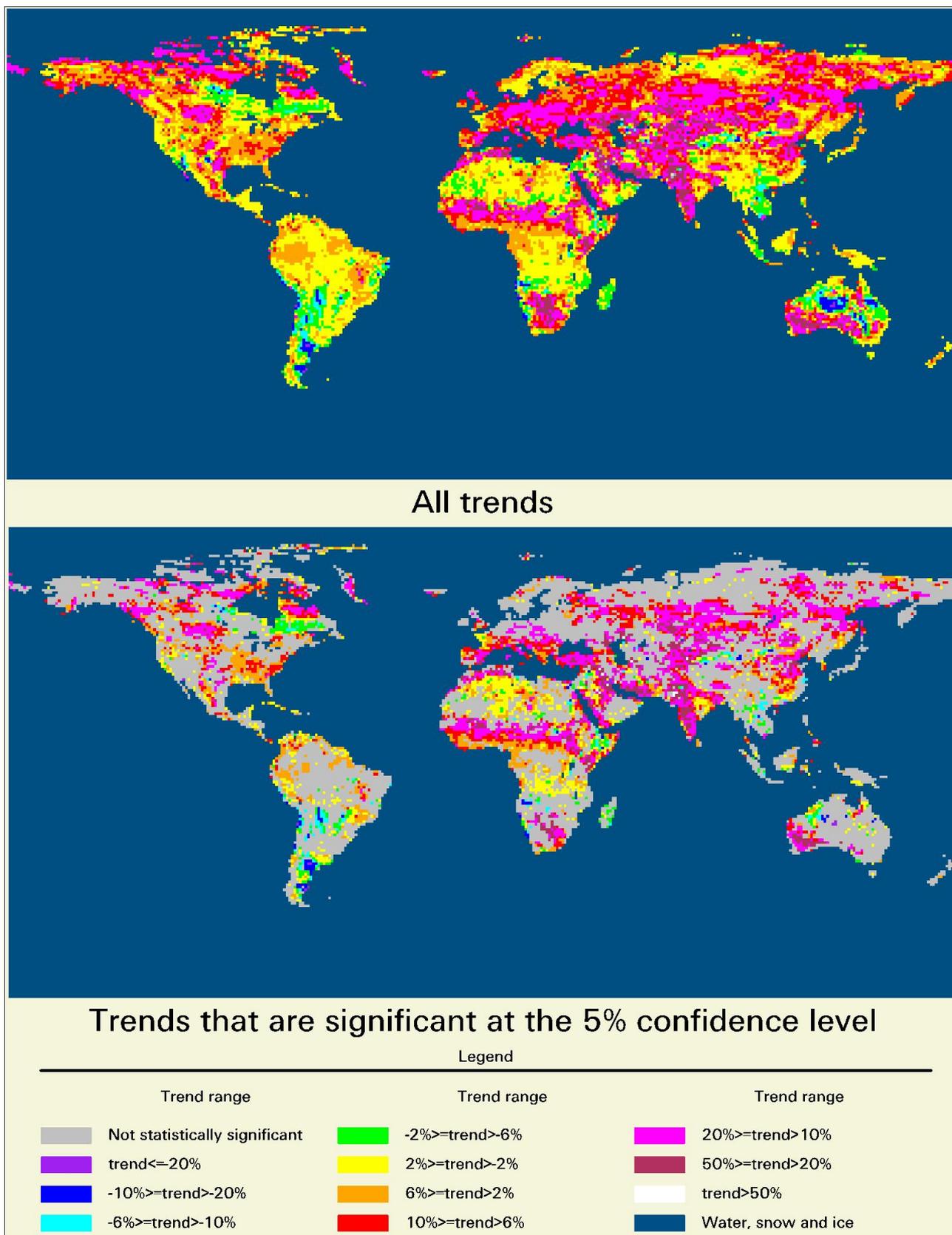


Figure 2: Percentage trends in NDVI as estimated from a quadratic least squares regression for the Fasir, version 2 dataset, for the period July 1982 – June 1998 (16 years). In the figure, (a) represents all trends, whilst (b) are only those trends that are significant at the 5% level as estimated using the Wald Test.

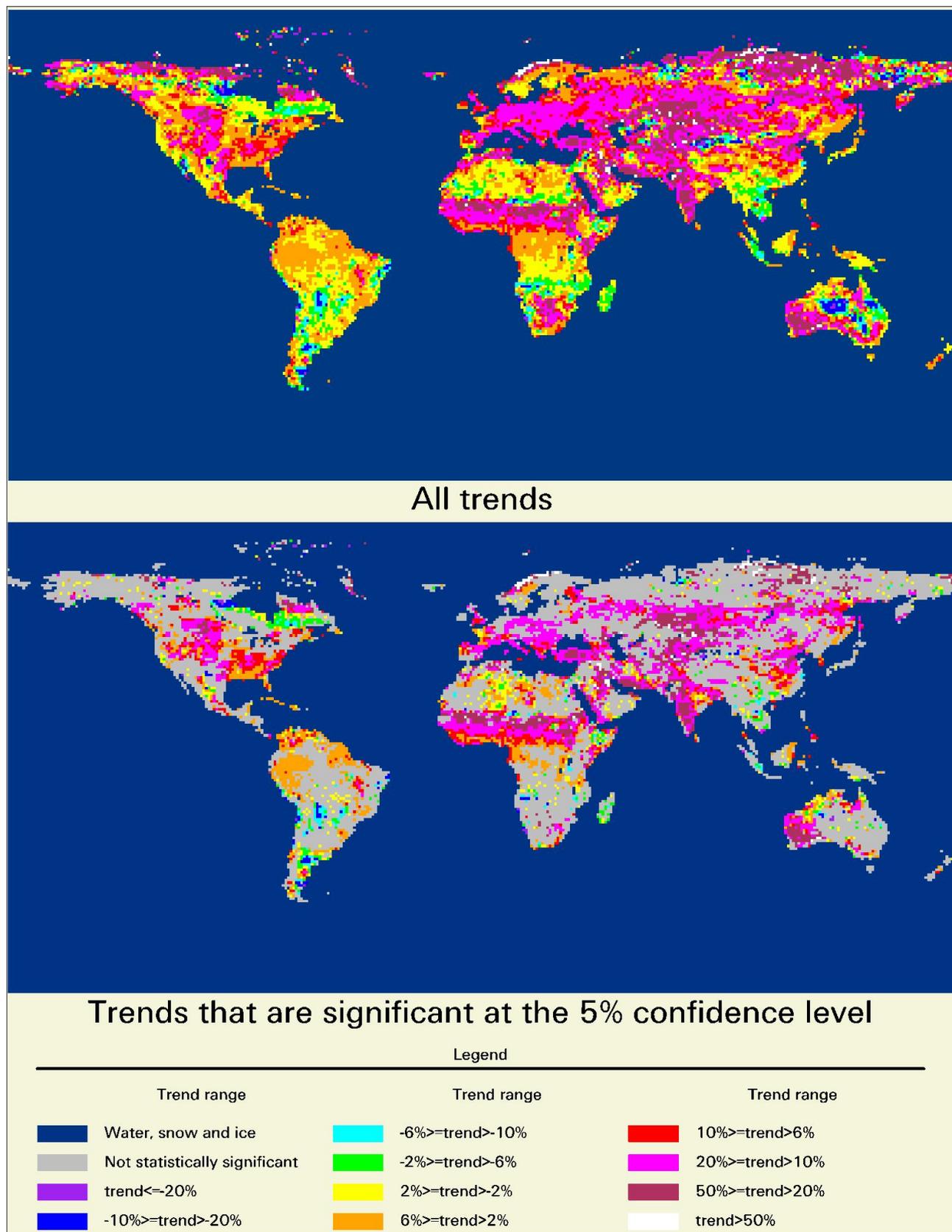


Figure 3: Percentage trends in NDVI as estimated from a quadratic least squares regression for the Fasir, version 4 dataset, for the period January 1982 – December 1999 (18 years). In the figure, (a) represents all trends, whilst (b) are only those trends that are significant at the 5% level as estimated using the Wald Test.

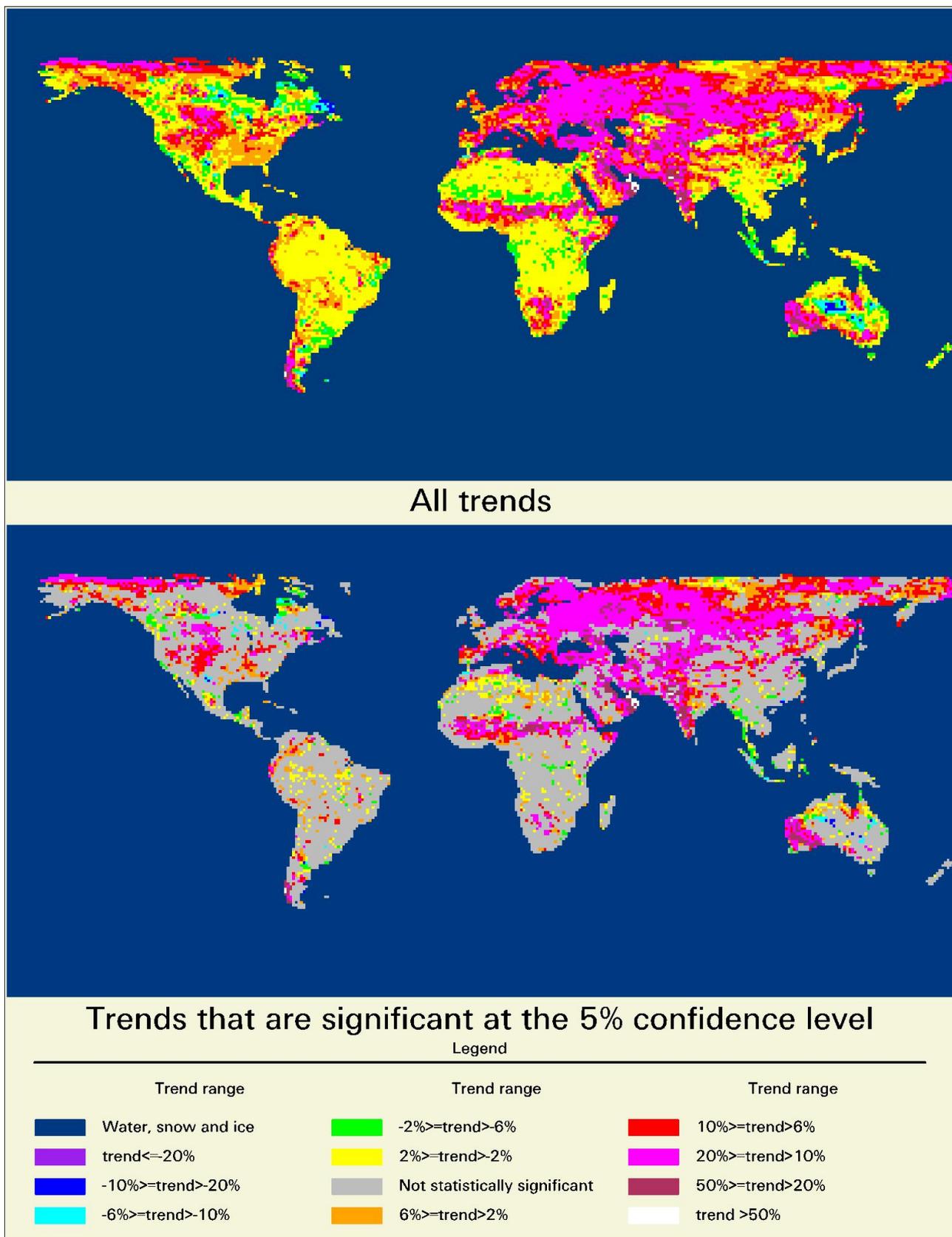


Figure 4: Percentage trends in NDVI as estimated from a quadratic least squares regression for the GIMMS, version 2 dataset, for the period July 1981 – June 2000 (19 years). In the figure, (a) represents all trends, whilst (b) are only those trends that are significant at the 5% level as estimated using the Wald Test.

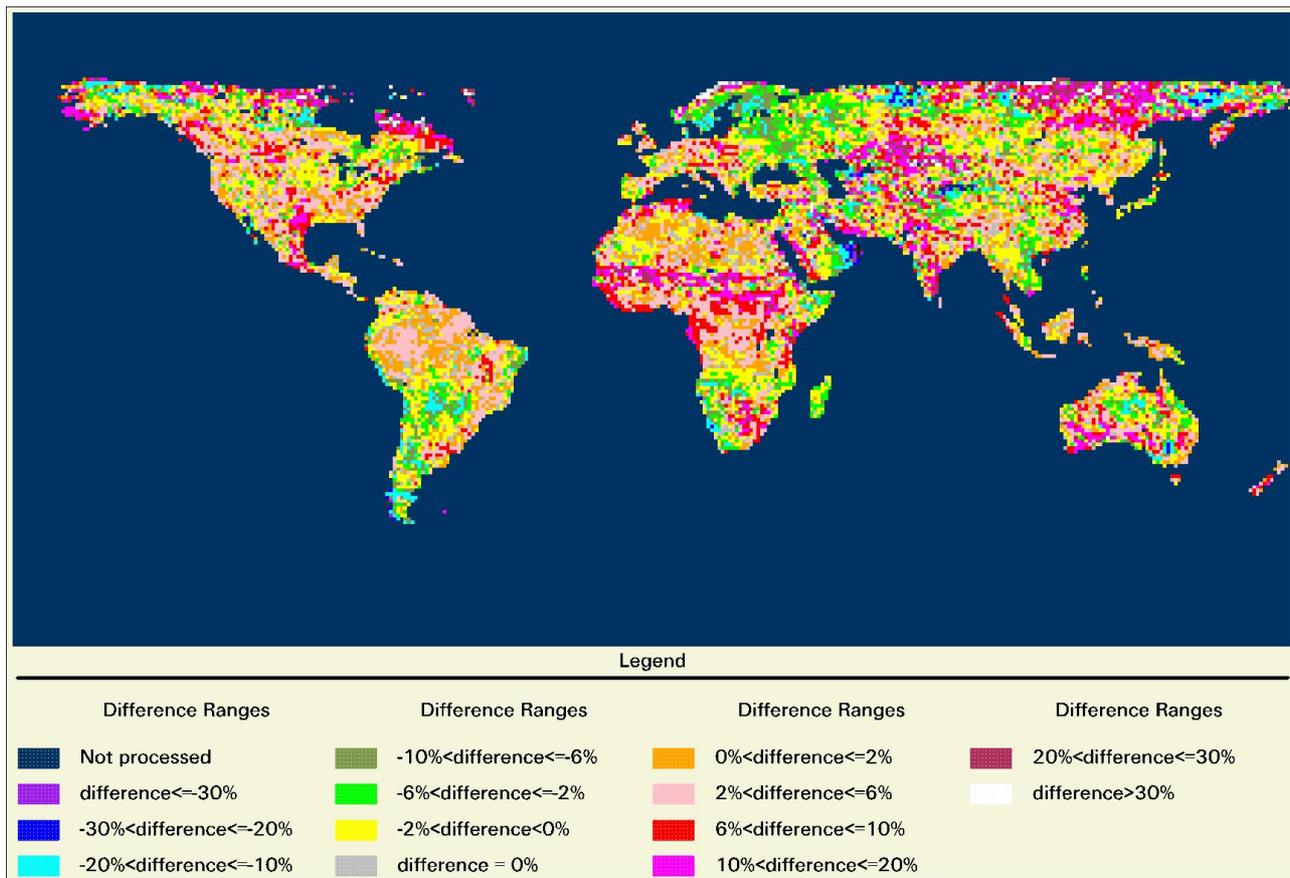


Figure 5: Difference in trends, Fasir V4 – GIMMS

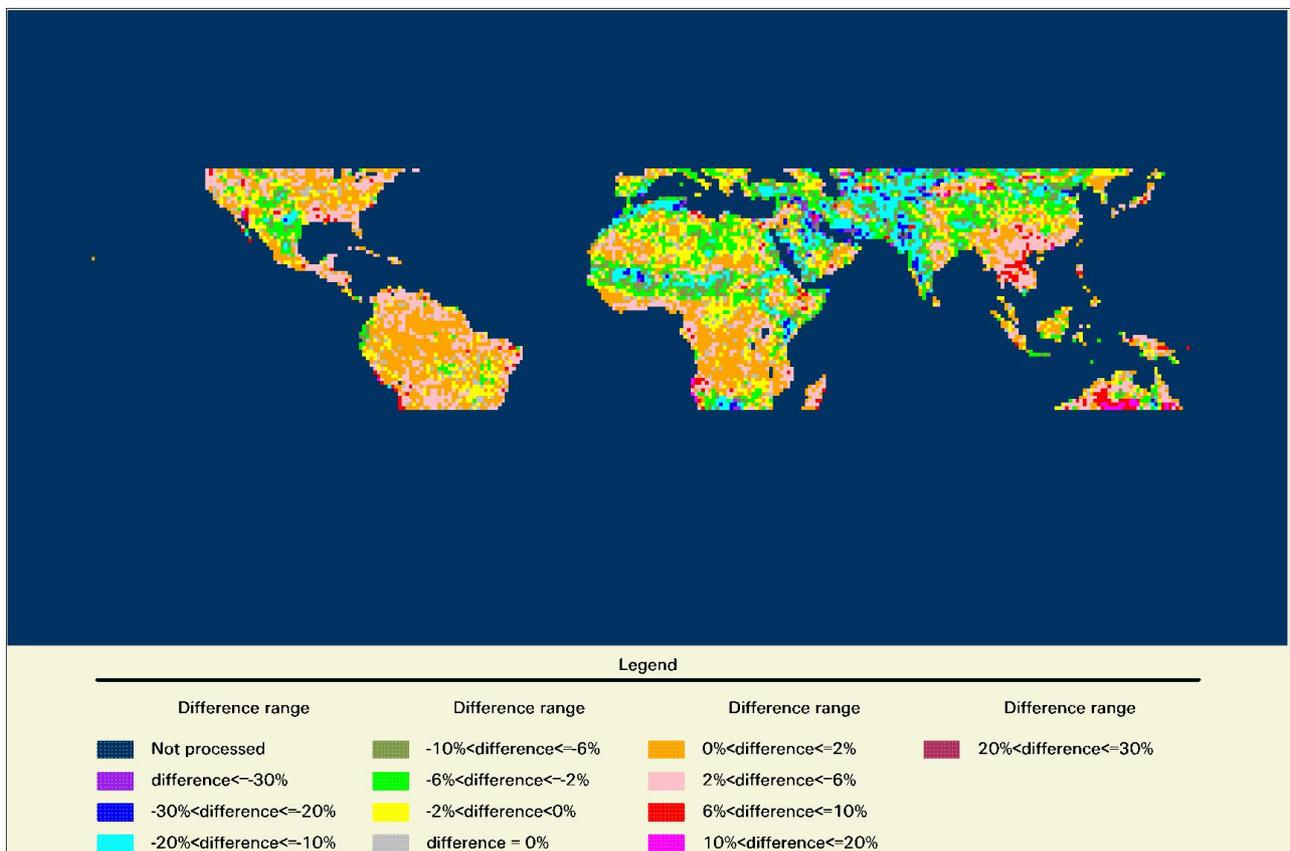


Figure 6: Difference in trends, Pathfinder - Fasir V2

Pairs of the trend images were used to derive difference images, depicted in Figures 4 and 5 for two of these differences. These difference images showed large differences where they were expected. It also showed, however, that there is a marked regionalisation that dominates the variations in the trends. This is particularly marked in the comparisons of the Fasir datasets with the GIMMS dataset (Figure 5), but it is still obvious in the comparisons between the two Fasir datasets and in the comparison of all the other datasets with the Pathfinder dataset (Figure 6).

Visual interpretation of the subtraction images suggests that there is some consistency in the regions identified in the different subtraction images. Some of these regions, such as the Sahel region, are located on the transition between the dry atmosphere of the Sahara, with higher aerosol loads to the north, and the moist atmosphere of tropical Africa. It is possible that the different processing steps are responding somewhat differently to differences in landcover in these regions, but it is more likely that the different processing methods respond somewhat differently to the atmospheric gradients across this linearity.

DISCUSSION ON THE COMPARISON OF THE LONG TERM TRENDS

To assess the likely validity of the results, the sources of error that may affect the calibration and correction of image data have been reviewed.

Fluctuations in atmospheric conditions

The atmosphere absorbs and scatters radiation due to the atmospheric chemicals, water droplets and dust. The atmospheric chemicals are relatively stable components of the atmosphere and hence their effects can be modelled and corrected. The moisture and dust content of the atmosphere are highly variable, both spatially and temporally. In addition, their effects interact, in that the moisture content of the atmosphere affects the absorption and scattering effects of the dust particles since the dust particles may absorb moisture and may tend to coagulate into larger particles. Whilst methods of correcting for these effects are evolving, they cannot be completely corrected at present, and historical data cannot be properly corrected for their effects. The effects of the atmospheric moisture vapour and aerosols are to reduce the contribution of the surface to the signal incident on the sensor, and hence they will reduce *NDVI*, where the reduction depends on the absorption and scattering that actually occurs. A positive trend in *NDVI* would require a decreasing trend in atmospheric moisture and/or aerosols. The trend of increasing surface temperature that has been found (23,24) should cause increasing evapotranspiration. If this vapour was converted into rain, then the associated conversion of the gas to the liquid should transfer energy to the atmosphere. However, no significant trend in tropospheric air temperature has yet been detected (25). It is more likely that the atmosphere is experiencing an increase in atmospheric moisture content than a reduction. Just as there are no data on which to analyse trends in atmospheric moisture content, similarly there is no long-term record of atmospheric aerosol content. The AERONET network of sun photometers (26), which started in about 1992, had established about 60 sites around the globe by 1998. Data from this network are showing that atmospheric aerosol loadings can fluctuate wildly with both time and position (27). However, the network has not had sufficient sites for long enough to conduct useful trend analysis over time. Current trends in atmospheric aerosols will be affected by reduced levels of pollution from western countries, such as reduced emissions from powerhouses and vehicles, but will be affected by increasing air pollution from other parts of the globe due to such things as increasing total power plant production and total vehicle emissions. It is also most likely that there is an increasing incidence of forest and grass fires as well as the increased emission of dust due to human activities. Whilst the sum of these effects is not known, it would seem unlikely that they will be contributing to a decreasing trend in atmospheric aerosols and pollutants, but their fluctuations may well contribute to regional fluctuations in the satellite signal. If air quality over Europe and the USA is improving, then the increases found over both continents in Figures 1-4 may be partly due to this cause. However, the increases detected over these continents are not in excess of those found for many other parts of the globe where air quality may be decreasing. There is thus no clear evidence in this data that the trends are due wholly or in part to the impact of man on atmospheric aerosol loads. If atmospheric aerosol loads are increasing, then it is likely that the trends detected in this study are conservative.

Variations in the sun-surface-sensor geometry

These variations occur with changes in the time of day that the image is taken, with the season, and with the distance that the object is away from the nadir point below the satellite. The longer the radiative path length, the greater the effect of the atmosphere on the signal thereby producing a reduction in *NDVI*. Trends towards later acquisition times may introduce a negative trend in *NDVI* since data acquisition occurs after solar noon, and variations in this geometry can introduce fluctuations in the signal.

Variations in surface reflectance

No surfaces reflect exactly equal amounts of energy in all directions, and the distribution varies between cover types and over time in each cover type, so that it is very difficult to predict patterns in the distribution. With vegetation the actual distribution varies with species, with canopy geometry and with the effect of the wind at the time of data acquisition. The integrative effect of the AVHRR sensor, with a 1.1 km field of view, means that the effect of these variations is relatively small near nadir from the sensor. However, the sensor can also acquire data up to 48° off the vertical axis, and at these larger angles the effect can become significant. The effect of these variations is to give a peak *NDVI* value at the optimum direction(s) and for *NDVI* to decrease more rapidly than the values in the two wavebands change away from these directions. This is unlikely to cause a trend in *NDVI* but may contribute to fluctuations in *NDVI* values.

Variations in the incident radiation at the top of the atmosphere

The solar radiation at the top of the atmosphere primarily varies during the season, due to the Earth's orbit and the tilt of the earth's axis. These variations can be corrected and so will not introduce a trend in *NDVI*.

Variations in the sensor

The variations in the sensor are physical, electronic and electro-optical. The physical component arises due to dust accumulating on the sensor optical surfaces and to any physical degradation that occurs to the optical or filter systems. The sensing devices have specific sensitivity characteristics at launch, and these degrade over time, reducing sensitivity to the incident radiation. All electronic systems contain internal current, and this can change over time. Sensor calibration corrections are derived prior to launch and have been shown to change over time. On the basis that the near infrared signal is much greater than the red signal from green vegetation, the largest effect of sensor degradation will be to reduce the near infrared signal more than the red signal, reducing *NDVI*. Any trend introduced due to sensor degradation will be reduced by the use of different sensors for different periods of time. With this time series, at least three sensors (NOAA-7, NOAA-9 and NOAA-11) have been used over the period of data acquisition.

Resampling of the data

Derivation of 1-degree spatial resolution, monthly data from 1.1 km resolution, daily data involves many resampling transformations. At each 1 km cell in the rectified and calibrated data, the highest *NDVI* value is selected in the month. A set of these cells is then used to compute the degree resolution data. The effect of this is to reduce the effects of the atmosphere on the derived image data. These processes are unlikely to introduce trends in the data, but can contribute to fluctuations.

This discussion and the results derived from the analysis suggest that the positive trend in *NDVI* is correct. However, the magnitude of this trend and its regional fluctuations are likely to be significantly affected by variations in atmospheric conditions since these have not yet been adequately dealt with in any of the datasets. In addition, the low radiation levels at high latitudes suggest that these effects may be more marked at these latitudes. Certainly the differences between the datasets are most marked at these latitudes.

THE COMPARATIVE ANALYSES OF THE DATASETS

To try and better understand the relationship between the datasets, it was decided to investigate the correlation between the datasets at each pixel. The method of analysis is based on that reported in (9) since this method provides the best fit linear regression for each set of twelve months of data and the Coefficient of Determination of that fit. This approach thus reduces the amount of data to be dealt with in the subsequent analysis. The method starts by assuming that the trajectory for the data in each twelve monthly period in each dataset is sinusoidal in form. In this comparison, since each dataset is taken from the same source data, the sinusoidal forms are assumed to vary in amplitude, mean value and phase shift, but not in frequency.

The independent dataset, $X = a_o + b_o \sin(c_o + d_o t)$ (2)

The dependent dataset, $Y = a_1 + b_1 \sin(c_1 + d_o t)$ (3)

The phase shift was dealt with by comparing the twelve monthly data in one dataset with the data in the same twelve monthly period in the other dataset, but conducting twelve regressions, each at a different monthly displacement between the datasets. Thus displacement = 0 means that January in dataset 1 is compared with January in dataset 2, etc, and displacement = 1 means that January in dataset 1 is compared with February in dataset 2 and so forth. Only the derived regression with the largest Coefficient of Determination or R^2 value was retained. In this way it can be found when there is a zero phase shift between the two datasets, that is $c_1 = c_o$ in Eqs. (2,3) and use this to remove t from the equations, to give:

$$Y = [a_1 - a_o(\beta_1)] + (\beta_1)X, \quad (4)$$

where $\beta_1 = (b_1/b_o)$, the ratio of the dependent to the independent amplitudes. As a consequence the data for each twelve monthly period in the two datasets were fitted by least squares to a line and the derived regression gain and offset, R^2 and phase shift values were retained. Since the Pathfinder data can contain anomalous values at high latitudes in winter, in both hemispheres, the method described in (9) was modified slightly to look for large residuals from the regression, and if this occurred, then the pair of data values (one from each dataset) that gave the outlier was rejected from the dataset, and the regressions conducted again.

The sine wave model provides an explanation of the derived regression parameters. Eq. (4) shows that the gain is the ratio of the amplitudes of the dependent to the independent dataset. The gain values should be close to unity for these datasets. If the gain value is close to unity, then the offset is the difference between the mean values. The offset should be close to zero in these comparisons.

For this analysis all of the datasets were converted to integer format, covering the same data range. To provide compatible data from each dataset, data was used from January 1982 until December 1998 and excluding 1995 for those comparisons that involved the Pathfinder dataset. The comparison of the Fasir 2 and the GIMMS dataset used data from July 1981 to June 2000.

Because the comparisons yield four parameters per annual cycle, this resulting data was further summarised by deriving mean and variance values for each of the four parameters. The work reported in this paper used the R^2 values derived from each year to set confidence and reliability limits, but it used the summarised results for the analysis of the gain values.

RESULTS OF THE COMPARATIVE ANALYSES

Statistics on the R^2 values for each comparison are given in Table 2. These R^2 values in each year were used to test the significance of the regression gain value, β_1 by the use of the F test at the 5% confidence level. The reliability was then assessed by the number of years within which the fit met this criterion, with two bounds being set, a 69% reliability of 11 or more years out of the 16 analysed (5/69 criterion), and 94%, or 15 or more years out of the 16 analysed (5/94 criterion). The results are shown in Figure 7 and summarised in Table 3. They show that there is a high level of

correlation between all of the datasets and that most areas have better than 94% reliability using the above definition. The reliability is highest between the Fasir 2 and the GIMMS datasets (Figure 7a) than between either and the Pathfinder dataset. The most significant discrepancy between the Fasir 2 and the GIMMS datasets is above 71°N where the best regression fit in each year does not meet the 5% confidence level for 69% of the years or more. Either or both of these datasets must be wrong in this part of the world. The lower reliability levels shown in the Amazon Basin, Arabia, Iran, New Guinea, central Asia and Australia, are in areas of small annual amplitude in NDVI. The most likely explanation for the poorer regression fits is that the variations between the datasets introduced in the processing are affecting the regression fitting when there is a small range in the data values used in the regressions.

Table 2: The Maximum, minimum and mean R² values found for each comparative analyses.

	Minimum	Maximum	Mean
Fasir2 and GIMMS	0.1450	0.9878	0.837
Pathfinder and Fasir2	0.0175	0.9729	0.753
GIMMS and Pathfinder	0.3198	0.9756	0.735

Table 3: Percentages of the land area for which the regression fit was significant at the 5% confidence level for 94% of the time (15 or more of 16 years), 69% of the time (11 to 14 of 16 years) and for less time.

	Met neither criteria	Met the 69% reliability criteria	Met the 94% reliability criteria
Fasir2 and GIMMS	1.72%	7.38%	90.90%
Pathfinder and Fasir2	1.58%	10.52%	87.90%
GIMMS and Pathfinder	1.99%	10.04%	83.97%

The gain and offset values were considered in the next stage of the analysis. In this study, only those areas that met the 5/94 rule described above were considered. In this analysis, the gain values should be in the vicinity of 1. They may vary somewhat from this as the data started in different formats and had to be transformed into the same range in the same format. Figure 9 summarises the gain settings derived from each comparison whilst Figure 8 depicts the regional variations.

Both the Fasir 2-Pathfinder (Pafe) and the Pathfinder – GIMMS (Page) comparisons show large areas of negative gain values in the high northern latitudes. The southern boundary of this region is most marked in the Pafe analysis, occurring at latitude 64°N.

Only in Scandinavia do the negative gain values occur south of this latitude in the Pafe analysis. This latitude is also the southern boundary of the negative gain values in the Page analysis in central Siberia, but the negative gain values in the Page analysis extend further south in Canada/Alaska, eastern and western Siberia. Given the conceptual basis for the gain values, negative values are most likely to occur when either there is negligible seasonality indicated in the data, or the corrections made to either or both datasets are creating very large differences between the datasets at different times during each annual cycle. Since the gain values being considered meet the 5/94 rule, they are highly consistent from year to year. This makes it unlikely that negligible seasonality will be the cause of the negative gain values, as could occur in the Amazon rainforest area. It is most likely that the data processing is introducing large differences into the datasets at different parts of the seasonal cycle.

Boundary zones of different gain values can be seen in the Fasir 2 – GIMMS (Fage) comparison in the Sahel, south of the Congo, along the Himalaya Mountains, along the Zagros Mountains in Iran, along the Sierra Madre in Mexico and in the vicinity of the Tien Shan and Altai mountains in west-

ern China. These linearities suggest that differences in processing are creating discontinuities between the datasets in these regions.

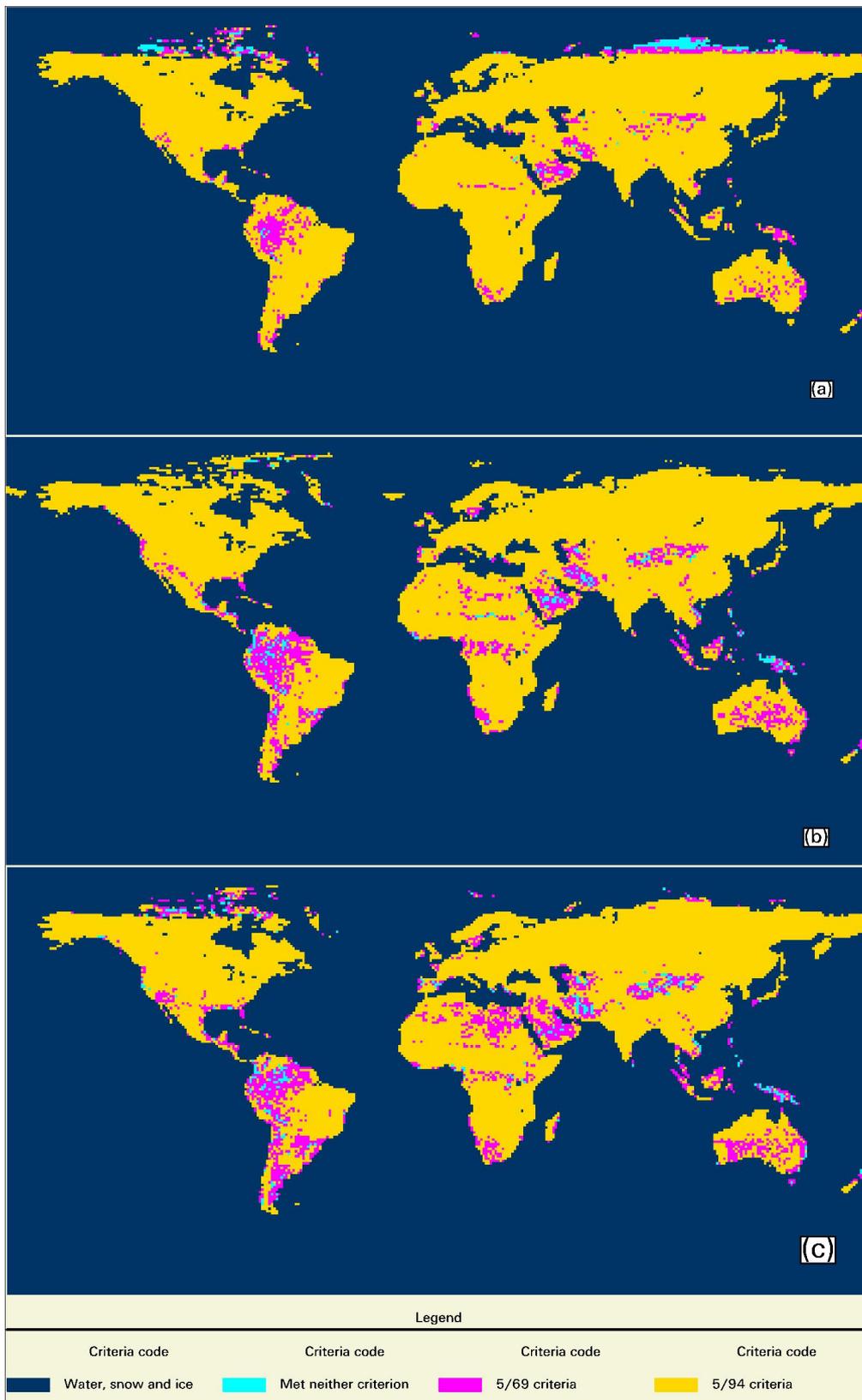


Figure 7: Areas that met the 5/94, 5/69 rules and those areas that met neither rule, representing the confidence level in the Regression and the reliability of the set of fits to the regression, (a) Fasir2 and GIMMS, (b) Pathfinder and Fasir 2 and (c) GIMMS and Pathfinder.

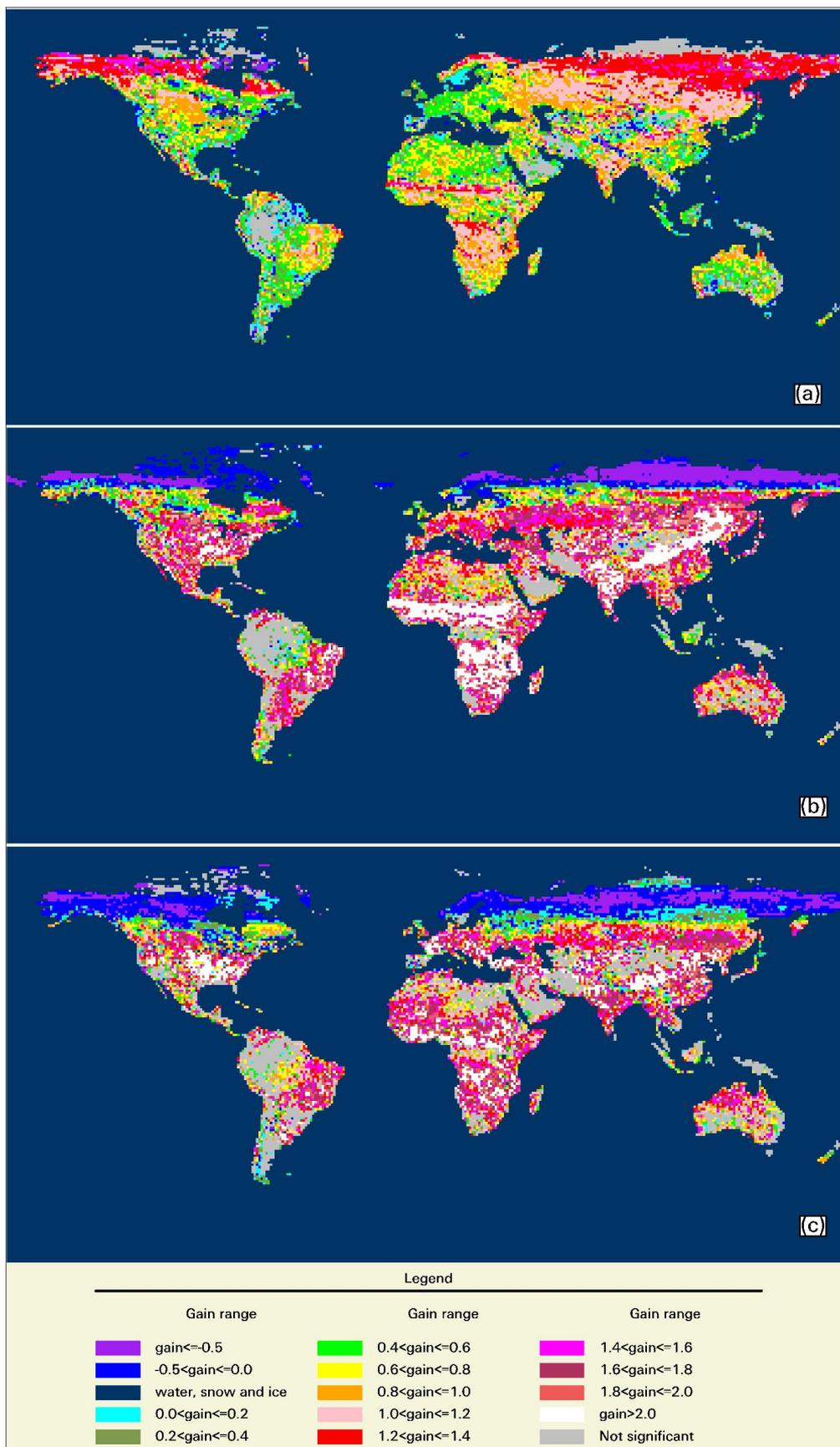


Figure 8: The gain settings for those areas that met the 95/94 rule as derived from each of the three comparisons, (a) Fasir2 and GIMMS, (b) Pathfinder and Fasir 2 and (c) GIMMS and Pathfinder.

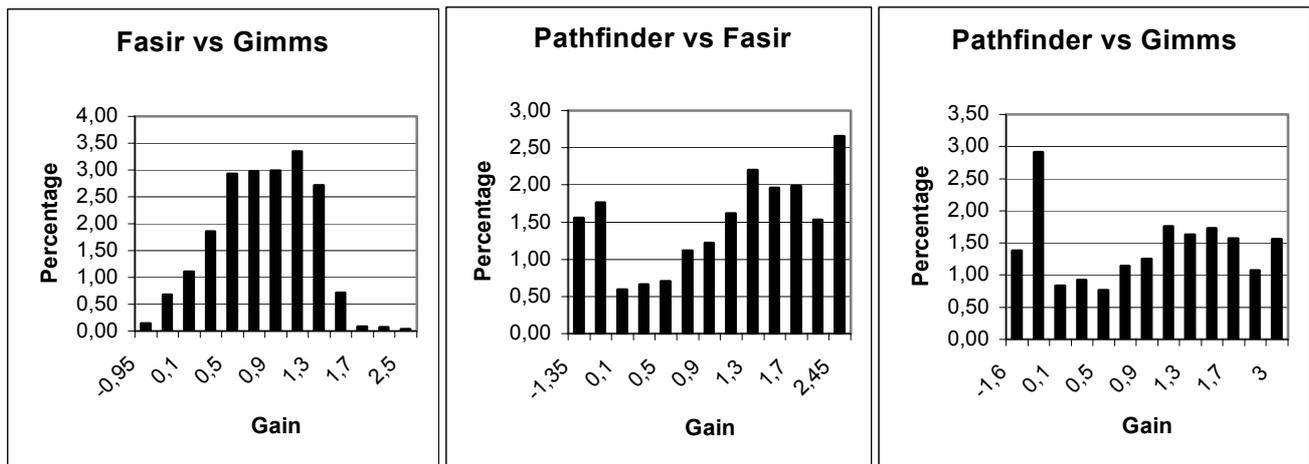


Figure 9: Histograms showing the distribution of the gain values derived from each comparison. A gain value of 1.0 indicates a one to one correspondence between the datasets, in terms of the gain values.

DISCUSSION

From this paper can be drawn the general conclusion that there has been a trend of increasing NDVI between 1981 and 2000, and that this trend is in the range 1.8 – 4.5% depending on the dataset used. The analysis shows the importance of correcting image data for atmospheric effects when the data are to be used to estimate surface parameter values. Without such a correction, the data can only be used to provide very generalised information.

That there is much more information in the data is obvious from the images. However that information will continue to be denied to us until the data are adequately corrected for the effects of the atmosphere. The analysis has shown that the Fasir and GIMMS datasets are more consistent between each other than either is with the Pathfinder dataset. Both the Fasir and the GIMMS datasets are designed to provide estimates of reflectance at the top of the atmosphere, whereas the Pathfinder data attempt to provide estimates of reflectance at the surface. For this reason alone, it is reasonable that there should be more consistency between the Fasir and the GIMMS datasets. But the fundamental need is for information on reflectance at the surface, since it is conditions on the surface that are fundamentally important for those who wish to use these datasets. This thus raises the question, “Are the Fasir and GIMMS datasets more consistent but less accurate than the Pathfinder dataset and the Pathfinder dataset less consistent but more accurate?”

It is not possible to answer this question in this analysis. However, it is the responsibility of each of the suppliers of the datasets to provide evidence as to the accuracy and reliability of their dataset. Whilst it is currently impossible to state the accuracy and reliability of historical datasets, it is not impossible to estimate their accuracy and reliability using current data. If the relationship between the image data and the effects of the atmosphere on that data remain the same, then current data for a period of twelve months can be used to provide a statement on the accuracy and reliability of the data. Whilst this is not as good as an assessment conducted during the acquisition of the image data itself, it is better than the current situation.

There is another reason why dataset suppliers should provide such information. Whilst the methods employed by each provider can be deduced from the documentation cited in this paper, none of this documentation covers the detail of the actual parameterisation that took place and which will have just as big an impact on the derived data as will the methods themselves. For example, which pixel values were actually chosen in the maximum value compositing, or how did the Libyan reflectance surface actually get used (The size of the area, the location, the method of registering the location, the method of estimating the surface reflectance, and so forth) in the calibration of the Pathfinder dataset. There are many similar subtleties embedded in each of the processing trains. It is impossible for users to backtrack through such detail, if only because there is no record of the

actual details for many of these steps. The only realistic solution is that the data suppliers provide evidence of the accuracy and reliability of their products.

A high priority for the satellite data providers should be to ensure the simultaneous acquisition of data suitable for atmospheric correction of the image data being collected, and the development of better methods of atmospheric correction of image data. Until this becomes routine, remote sensing will not be able to fully exploit the data being acquired to derive information on the biophysical processes that are taking place on the earth's surface.

ACKNOWLEDGEMENTS

The authors acknowledge the provision of the datasets by the suppliers as described in the paper.

The work was conducted as part of the "Temporal Change and Image Segmentation project" funded in part by the Danish ESA Following Research Fund.

The very constructive comments of the reviewers are gratefully acknowledged.

REFERENCES

- 1 Tucker C J, H E Dregne & W W Newcombe, 1991. Expansion and contraction of the Saharan desert from 1980 to 1990. Science, 253: 299-301
- 2 Goward S N, A Kerber, D G Dye & V Kalb, 1987. Comparison of North and South American biomes from AVHRR observations. Geocarto, 2(1), 27-40
- 3 Myneni R B, C D Keeling, C J Tucker, G Asrar & R R Nemani, 1997. Increased plant growth in the northern high latitudes from 1981-1991. Nature, 386: 698-701
- 4 Tucker C J, D A Slayback, J E Pinzon, S O Los, R B Myneni & M G Taylor, 2001. Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. International Journal of Biometeorology, 45(4): 184-190
- 5 Los S O, G J Collatz, L Bounoua, P J Sellers & C J Tucker, 2001. Global interannual variations in sea surface temperature and land surface vegetation, air temperature, and precipitation. Journal of Climate, 4 (7): 1535-1549
- 6 Tucker C J & S E Nicholson, 1999. Variations in the Size of the Sahara Desert from 1980-1997. Ambio, 28: 587-591
- 7 De Fries R, M Hansen & J Townshend, 1995. Global discrimination of landcover types from metrics derived from AVHRR pathfinder data. Remote Sensing of Environment, 54: 209-222
- 8 Loveland T R, B C Reed, J F Brown, D O Ohlen, Z Zhu, L Yang & J W Marchant, 2000. Development of a global landcover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21(6): 1303-1330
- 9 McCloy K R & W Lucht, 2004. Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model. IEEE Transactions on Geoscience and Remote Sensing, 42(1): 140-153
- 10 http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/FTP_SITE/readmes/pal.html
- 11 Rao C R N, 1993. Degradation of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-9 spacecraft: Assessment and recommendations for corrections. NOAA Technical Report NESDIS-70 (NOAA/NESDIS, Washington, DC) 25 pp.
- 12 Gordon H R, J .W Brown & R H Evans, 1988. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone colour scanner. Applied Optics, 27:2111-2122

- 13 http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/ISLSCP/DATASET_DOCUMENTS/FSR_NDVI.html
- 14 Sellers P J, S O Los, C J Tucker, C O Justice, D A Dazlich, G J Collatz & D A Randall, 1994. A global 1×1 degree NDVI dataset for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. International Journal of Remote Sensing, 15(17): 3519-3545
- 15 Los S O, G J Collatz, P J Sellers, C M Malmström, N H Pollack, R S DeFries, L Bounoua, M T Parris, C J Tucker & D A Dazlich, 2000. A global 9-year biophysical land-surface data set from NOAA AVHRR data. Journal of Hydrometeorology, 1: 183-199
- 16 Sellers P J, S O Los, C J Tucker, C O Justice, D A Dazlich, G J Collatz & D A Randall, 1996. A revised land surface parameterization (SiB-2) for atmospheric GCMs. Part 2: The generation of global fields of terrestrial biophysical parameters from satellite data. Journal of Climate, 9: 706-737
- 17 Los S O, C O Justice & C J Tucker, 1994. A global 1 by 1 degree NDVI dataset for climate studies derived from the GIMMS continental NDVI data. International Journal of Remote Sensing, 15(17): 3493-3518
- 18 <http://ftpwww.gsfc.nasa.gov/gimms/htdocs/>
- 19 Kidwell K B, 1988. NOAA polar orbiter data (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10 and NOAA 11) users guide. (NOAA, Washington, DC)
- 20 Los S O, 1993. Calibration adjustment of the NOAA AVHRR Normalized Difference Vegetation Index without recourse to component channel 1 and 2 data. International Journal of Remote Sensing, 14: 1907-1917
- 21 McCulloch C E & S R Searle, 2001. Generalised, linear and mixed models. Wiley Series on Probability and Statistics (John Wiley and Sons, New York) 358 pp.
- 22 Loveland T R & A S Belward, 1997. The IGBP-DIS global 1-km landcover dataset DISCover: first results. International Journal of Remote Sensing, 18: 3289-3295
- 23 <http://www.giss.nasa.gov/data/update/gistemp/>
- 24 <http://www.grida.no/climate/vital/17.htm>
- 25 <http://cdiac.esd.ornl.gov/trends/temp/sterin/sterin.html>
- 26 Holben B N, T F Eck, I Slutsker, D Tanre, J P Buis, A Setzer, E Vermote, J A Reagan, Y J Kaufman, J Nakajima, F Lavenue, I Jankowiak, & A Smirnov, 1998. Aeronet – A federated network and data archive for aerosol characterisation. Remote Sensing of Environment, 66: 1-16
- 27 <http://www.gisdevelopment.net/aars/acrs/1999/ps6/ps61220.shtml>