

Filtering Pathfinder AVHRR Land NDVI data for Australia

J.L. Lovell, R.D. Graetz

CSIRO Earth Observation Centre
GPO Box 3023
Canberra, ACT, 2601
Australia

Contact

Voice: +61-2-6216 7200
Fax: +61-2-6216 7222
Jenny.Lovell@eoc.csiro.au
Dean.Graetz@eoc.csiro.au

Abstract

The Pathfinder AVHRR Land (PAL) data set is a 13 year time series of consistently processed data which can be used in modelling and monitoring of climate and land cover change. The data are derived from AVHRR global area coverage data, resampled to 8km resolution and 10-day composite images. The compositing reduces the effect of cloud, but there is still significant cloud contamination and other noise present in the PAL dataset. The best index slope extraction algorithm was tested and a modified algorithm applied to the PAL dataset for the Australian continent. This report describes the modifications and presents some examples of the reduction in noise which is achieved by this method. The filtered dataset is available for use within CSIRO.

1 Introduction

The NASA/NOAA-sponsored Earth Observing System Pathfinder projects produced time series of global data sets for use in climate change research. The Pathfinder AVHRR Land (PAL) dataset comprises global, 10-day composite images of AVHRR data covering the period mid-July 1981 to September 1994 (James and Kalluri 1994). This report describes an NDVI filtering process which was applied to the PAL data for the Australian continent, to minimise noise such as cloud contamination.

2 PAL processing

The PAL dataset was derived from global area coverage data (GAC) acquired by the 5-channel AVHRR instruments, which is produced by onboard averaging and sampling to a resolution of 4km at nadir. The GAC data were navigated and corrections for sensor degradation, Rayleigh scattering and ozone absorption were applied. Cloud contamination was flagged by the CLAVR (Stowe *et al.* 1991) process based on 2×2 pixel arrays. NDVI was then calculated from the channel 1 and 2 reflectances. Re-sampling to 8km resolution was performed by forward, nearest neighbour mapping, selecting the highest NDVI value from all pixels within 42° of nadir. Small gaps were filled by copying values from adjacent pixels. Pixels for which the solar zenith angle was greater than 80° were not used. These re-sampled data were composited in quasi 10-day periods according to the maximum value of NDVI. The CLAVR cloud flags were not used in the re-sampling or compositing processes.

The PAL data are mapped to the Goodes Interrupted Homolosine projection. The data layers and their units are listed in Table 1.

Table 1: Pathfinder AVHRR Land data set layers.

Data	Units
NDVI	ndvi
CLAVR flag	Numeric value from lookup tables
QC flag	Numeric value from lookup tables
Scan angle	radians
Solar zenith angle	radians
Relative azimuth	radians
Ch 1 reflectance	% reflectance
Ch 2 reflectance	% reflectance
Ch 3 btemp	deg K
Ch 4 btemp	deg K
Ch 5 btemp	deg K
Day of year	DDD.HH

Compositing by maximum value of NDVI is used to minimise the effect of cloud in NDVI data, the assumption being that all contamination results in lower values of NDVI. This can be effective if the compositing period is quite long (2–4 weeks). However if the period is too long, short-term changes in vegetation condition may be lost. The PAL 10-day compositing period was not sufficient to remove all cloud affected pixels. Gutman and Ignatov (1996) found that some improvement could be made by taking the CLAVR flags into account and reducing the dataset to a coarser spatial or temporal resolution. However, they note that a large number of pixels are labelled as ‘mixed’. They maintain that some of these are in fact cloud-free and have been wrongly classified by the CLAVR algorithm. Prince and Goward (1996) also noted the large number of flagged pixels, commenting that in some cases as many as 90% of pixels in a 10-day composite image are flagged with some type of cloud cover.

3 NDVI noise removal techniques

Maximum value composition (MVC) is often used to minimise the the effects of cloud contamination in NDVI data but, as already noted, this technique has inherent problems. The length of the compositing period affects the success of the noise removal. A long period will be more effective in

removing noise, but may result in the loss of important short term changes. Spurious high values due to data transmission errors will not be removed by MVC.

The best index slope extraction (BISE) method of Viovy *et al.* (1992) was proposed as an alternative to MVC. This method is based on the predictability of seasonal changes in vegetation and accounts for the fact that growth and senescence phases are often asymmetric. It can also accommodate sudden changes due to fire, deforestation or crop harvest.

Sudden rises and falls in NDVI values are not compatible with the gradual process of regrowth, but are a feature of changing cloud conditions or viewing angles. This is the basis for the BISE algorithm, which applies the following rules. From the first date of the time series, the algorithm searches forward, accepting points if their value is higher than the preceeding one. If a decrease is found, the search is continued over a pre-defined sliding period (e.g. 30 days). The decrease is only accepted if there are no points within the sliding period with values greater than the previous high value minus 20% of the size of the decrease. That is, the decrease is only accepted if it is followed by a gradual increase consistent with regrowth. If an acceptable point is found within the sliding window, this point is chosen and the low point is rejected. Intermediate values are filled by interpolation and the search begins again from the last accepted point. The threshold of 20% of the decrease was chosen empirically for West African conditions. It should be adjusted to suit the typical regrowth rate of the area under consideration. Similarly, the length of the sliding period should be adjusted to suit the predominant cloud conditions. In addition to the tests applied within the sliding window, the BISE algorithm rejects any points with a random increase of greater than 0.1 as such fluctuations from one day to the next are unlikely from natural surfaces and are more probably the result of data transmission errors.

4 Modified BISE filtering of PAL NDVI

The BISE algorithm is designed for use with daily NDVI data. The performance is not expected to be as good with the PAL 10-day composites as artifacts of the compositing process are already present in this dataset. However, the PAL data contains some noise of the type that could be removed by a BISE-type procedure. Adjustments have been made to the BISE rules so that the algorithm works more effectively with the PAL NDVI data for Australia.

The test for spurious high points was modified to look for a spike (i.e. an increase immediately followed by a decrease). This refinement was necessary because the time between consecutive points is not equal in a composited dataset. The time may range from 1 to 20 days, depending on which dates within the compositing decads were chosen, so a simple threshold test is not appropriate.

The BISE algorithm performed poorly when there was a long-term, gradual decrease in NDVI. The test within the sliding window was modified to take into account the local gradient of the data. The local gradient was calculated for each point in the time series, over the same time period of the sliding window. The resulting time series of gradients was then smoothed. When the gradient was negative, the test value for high points within a sliding window was calculated by extrapolating from the high point before the decrease according to the local gradient, then subtracting 20% of the decrease.

Figures 1 and 2 show the results of the modified BISE filtering on four test sites using sliding windows of 3 decads (figure 1) and 6 decads (figure 2). The threshold for spurious increases was set at 0.1. The sites were chosen for their different vegetation types and climatic zones. The first is an area of forest in Western Australia. This site may be expected to be cloudy quite often due to its proximity to the coast. The second site is an area of wheat crop in Western Australia and clearly shows the annual growth and harvest cycle. The third time series is from the Nullabor plain, an arid and largely cloud-free region. The final test site was in tropical forest in Queensland. This area experiences long periods of cloud cover, particularly during the wet season.

The results shown in figures 1 and 2 show that the modified BISE filtering greatly reduces the noise in the NDVI time series. The filtered data are realistic and maintain seasonal variation. The

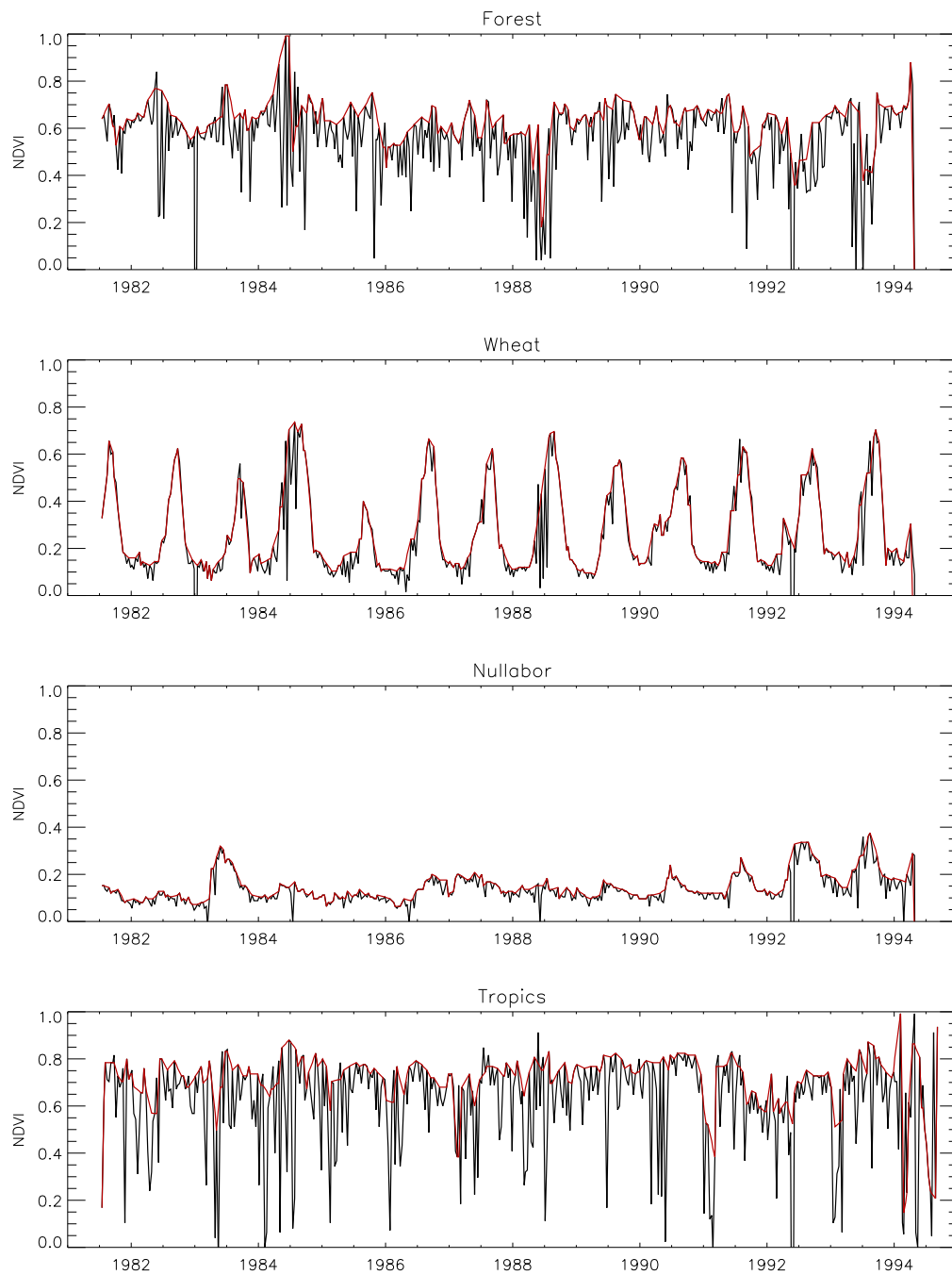


Figure 1: NDVI from Pathfinder AVHRR land dataset for four test pixels through the time series July 1981 to September 1994 (black lines) filtered with a window of 3 decades and spike threshold of 0.1 (red lines).

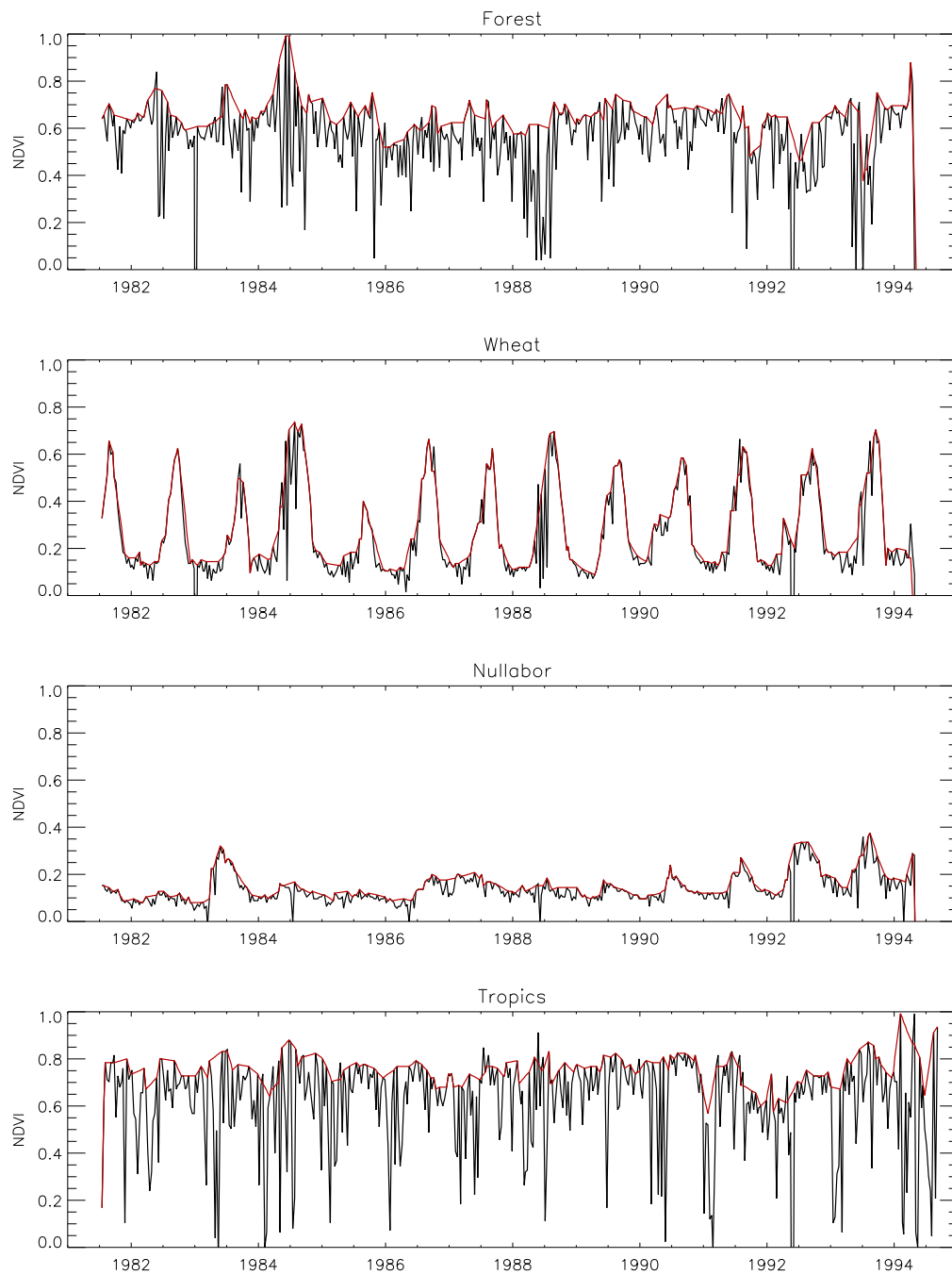


Figure 2: NDVI from Pathfinder AVHRR land dataset for four test pixels through the time series July 1981 to September 1994 (black lines) filtered with a window of 6 decades and spike threshold of 0.1 (red lines).



Figure 3: NDVI from Pathfinder AVHRR land dataset for the compositing period 21–31 August 1981, in northern Western Australia. Left original, right filtered.

sliding window in the first example (3 compositing periods) was not long enough to remove all the cloud effects in the forest and tropics sample sites, but is reasonable for the less cloudy wheat crop and Nullabor sites. The 6 decad window is more effective on the cloudy sites and still retains most of the sharp changes at the other sites. Ideally, the length of the window should be varied across the continent in different climatic zones, but a window length of 6 compositing periods was judged to be a good compromise and was applied for all of Australia.

Figures 3–7 are examples of NDVI images of Australia before and after the modified BISE filtering. The colours range from red (low NDVI, dry vegetation and bare soil) through yellow to green (high NDVI, green vegetation). The NDVI images were reprojected from Goodes Interrupted Homolosine to Plate Carée prior to filtering. These images demonstrate a variety of noise effects which were removed or reduced by the modified BISE.

Figure 3 is an area of northern Western Australia. A line of spurious values was removed by the filtering.

Figure 4 shows evidence of cloud (mottled red patches) in south-eastern Australia during May 1982. The filtered image is more spatially consistent and accords with the expected greenness of this area. Another example of cloud removal is shown in figure 5. This image shows the far north of Australia during the tropical wet season in February 1992. There are large red patches in the original data. The result of the filtering is again a much more spatially consistent image. Other similar examples can be found throughout the dataset.

Both figures 6 and 7 show the southern part of Australia and illustrate large areas of missing data. This is due to rejection of data when the solar zenith angle was greater than 80° . Some of these missing areas were repaired (interpolated) by the filtering process. Sharp discontinuities between adjacent satellite passes evident in the south east of figure 6 and in southern Western Australia in figure 7 were also corrected by the filtering.

These examples demonstrate that, although the noise removal was by temporal filtering performed on individual pixel time series, the effect is to improve the spatial uniformity of the images. Another test of the process is to look for seasonal changes expected in a series of images. Figure 8 shows images for the middle decad (days 11–20) of each month in 1987.

The evergreen forest regions in southern Western Australia, Tasmania and along the eastern sea-board are green throughout this time series. Some small red patches can be seen in Tasmania during June and July. This is most likely due to residual cloud and suggests that the length of the filtering window was not sufficient to remove persistent winter cloud in this area. Southern grasslands, shrublands and croplands are greenest in spring. This is evident in the extensive green area in south-eastern Australia which is largest in September. The tropical regions are greenest in March and April following the wet season. A seasonal crop cycle is most clearly illustrated by the wheat growing region of southern Western Australia. This is seen as a red patch from January to May and again in November and December. The red area becomes green in the late winter-spring months when the crop is growing vigorously. A similar trend can also be seen in smaller cropping areas of south-eastern Australia.

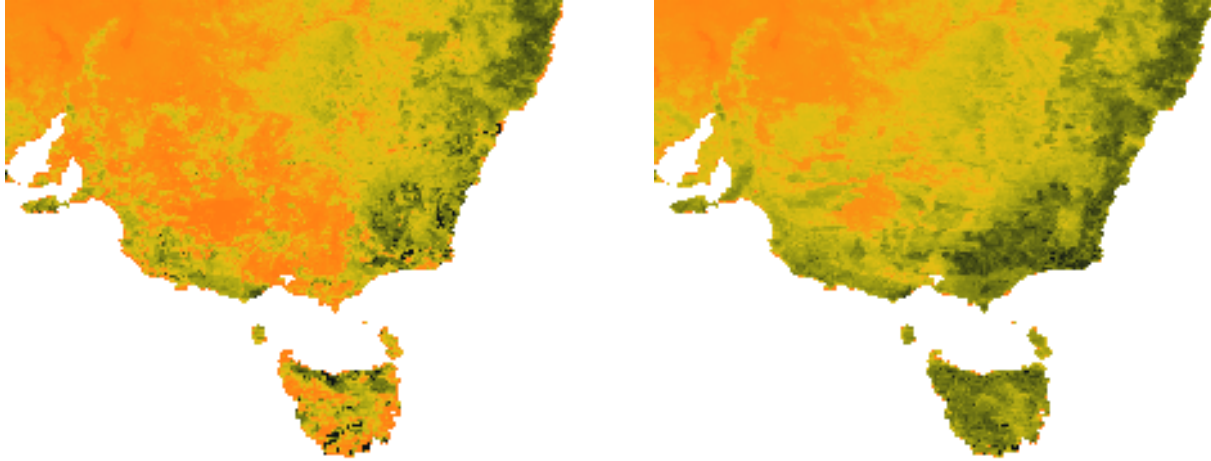


Figure 4: NDVI from Pathfinder AVHRR land dataset for the compositing period 21–31 May 1982, in south eastern Australia. Left original, right filtered.

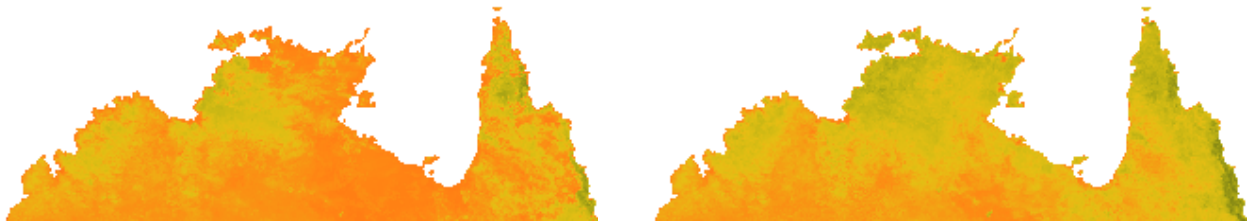


Figure 5: NDVI from Pathfinder AVHRR land dataset for the compositing period 21–28 February 1992, in northern Australia. Left original, right filtered.

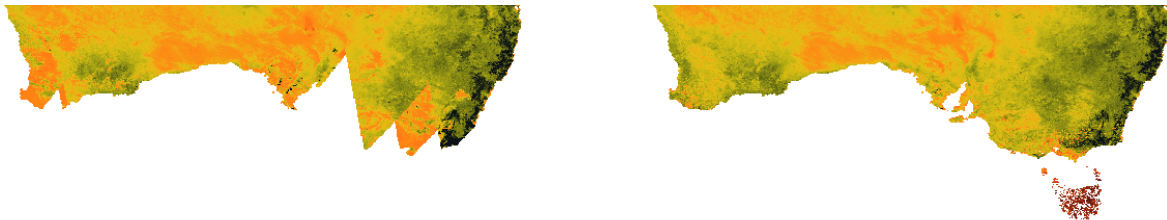


Figure 6: NDVI from Pathfinder AVHRR land dataset for the compositing period 11–20 June 1988, in southern Australia. Left original, right filtered.

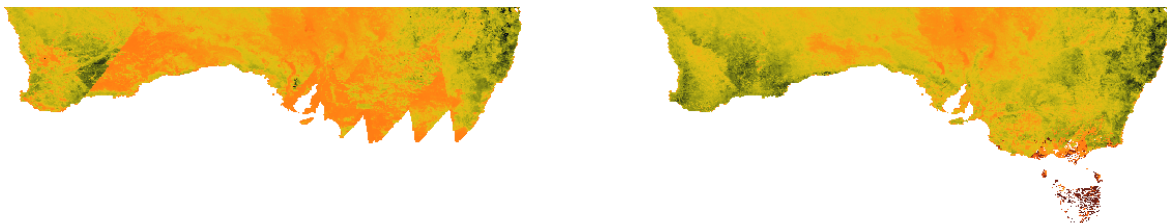


Figure 7: NDVI from Pathfinder AVHRR land dataset for the compositing period 21–30 June 1993, in southern Australia. Left original, right filtered.

All these observations of seasonal effects are realistic and show a smooth progression through the year. Thus the temporal filter has preserved seasonal changes.

The product of the modified BISE filtering is useful but not definitive. There are flaws still present and users must be aware of this. Even so, it is hoped that this dataset will be of value to its users.

EOC staff are in the process of producing a consistent AVHRR time series (including NDVI) for the Australian continent of the highest quality. This will employ CSIRO best practice algorithms for calibration, navigation, atmospheric correction, cloud clearing and removal of the effects of different solar and viewing angles (BRDF).

5 Data availability

The filtered dataset as described in this report is available for use within CSIRO. An extract of the readme file distributed with the dataset is shown in the appendix. The dataset can be obtained by contacting the authors.



Figure 8: NDVI time series for 1987.

Appendix

aus81_94ndvi_bise.readme

Edition 20000524_JLL

The above data set has been derived from the NASA NOAA Pathfinder AVHRR Land (PAL) data set (edition 2) by staff of the CSIRO Earth Observation Centre (EOC).

The data set consists of one image containing 474 channels of NDVI in byte format (DN).

The image format details are 450 lines by 600 pixels.

The geocoding is ULC (107.9574, -9.4650) and LRC (155.8774, -45.3850).

The conversion from DN (0-255) to NDVI (-1.0 - 1.0) is:

$(DN-128)*0.008$

Each channel represents a 10 day (decad) compositing period as described in the PAL Users Manual.

Channel 1 represents compositing period 20: JUL_11-20. Thereafter the channels represent the sequence of the remaining 36 compositing periods of each year, 1981-1987.

A table relating compositing period to dates is included below. A second table assigns the channel numbers to decimal years for plotting.

The EOC value-adding included:

(1) Reprojecting the original PAL data set from Goodes Interrupted Homolosine to Plate Caree (Geographic Lat/Long) with a spatial resolution (pixel size) of 0.08 deg., or approx. 8 km;

(2) Filtering the time series of each pixel to remove 'spikes', clouds and cloud shadows using a modified version of

the Best Index Slope Extraction (BISE) algorithm.
This algorithm used a search window of 6
decads and a NDVI change threshold of 0.1 per decad.

This data set is provided as a service
to those researchers who require a time
series for exploration and model
development.

This product is useful but not
definitive. The flaws still resident
are well known to EOC staff and the user
must accept this. Even so, we hope that
this data set will be of value to
its users.

EOC staff are in the process of
producing a consistent NDVI time series
of the highest quality.

Please direct all feedback to:

Susan.Campbell@eoc.csiro.au
Jenny.Lovell@eoc.csiro.au
Dean.Graetz@eoc.csiro.au

#,Comperiod,Start,Finish,Duration,Mid date (DoY),DoY as fraction of year.
1,JAN_01-10,1-Jan,10-Jan,10,5,0.01
2,JAN_11-20,11-Jan,20-Jan,10,15,0.04
3,JAN_21-31,21-Jan,31-Jan,11,26,0.07
4,FEB_01-10,1-Feb,10-Feb,10,36,0.10
5,FEB_11-20,11-Feb,20-Feb,10,46,0.13
6,FEB_21-28*,21-Feb,28-Feb,8,54,0.15
7,MAR_01-10,1-Mar,10-Mar,10,64,0.18
8,MAR_11-20,11-Mar,20-Mar,10,74,0.20
9,MAR_21-31,21-Mar,31-Mar,11,85,0.23
10,APR_01-10,1-Apr,10-Apr,10,95,0.26
11,APR_11-20,11-Apr,20-Apr,10,105,0.29
12,APR_21-30,21-Apr,30-Apr,10,115,0.32
13,MAY_01-10,1-May,10-May,10,125,0.34
14,MAY_11-20,11-May,20-May,10,135,0.37
15,MAY_21-31,21-May,31-May,11,146,0.40
16,JUN_01-10,1-Jun,10-Jun,10,156,0.43
17,JUN_11-20,11-Jun,20-Jun,10,166,0.45
18,JUN_21-30,21-Jun,30-Jun,10,176,0.48
19,JUL_01-10,1-Jul,10-Jul,10,186,0.51
20,JUL_11-20,11-Jul,20-Jul,10,196,0.54
21,JUL_21-31,21-Jul,31-Jul,11,207,0.57
22,AUG_01-10,1-Aug,10-Aug,10,217,0.59

23,AUG_11-20,11-Aug,20-Aug,10,227,0.62
 24,AUG_21-31,21-Aug,31-Aug,11,238,0.65
 25,SEP_01-10,1-Sep,10-Sep,10,248,0.68
 26,SEP_11-20,11-Sep,20-Sep,10,258,0.71
 27,SEP_21-30,21-Sep,30-Sep,10,268,0.73
 28,OCT_01-10,1-Oct,10-Oct,10,278,0.76
 29,OCT_11-20,11-Oct,20-Oct,10,288,0.79
 30,OCT_21-31,21-Oct,31-Oct,11,299,0.82
 31,NOV_01-10,1-Nov,10-Nov,10,309,0.85
 32,NOV_11-20,11-Nov,20-Nov,10,319,0.87
 33,NOV_21-30,21-Nov,30-Nov,10,329,0.90
 34,DEC_01-10,1-Dec,10-Dec,10,339,0.93
 35,DEC_11-20,11-Dec,20-Dec,10,349,0.96
 36,DEC_21-31,21-Dec,31-Dec,11,360,0.99

#,Year,Comperiod,Decimal Year+DoY,Decimal DoY

1,1981,20,1981.54,0.54
 2,1981,21,1981.57,0.57
 3,1981,22,1981.59,0.59
 4,1981,23,1981.62,0.62
 5,1981,24,1981.65,0.65
 6,1981,25,1981.68,0.68
 7,1981,26,1981.71,0.71
 8,1981,27,1981.73,0.73
 9,1981,28,1981.76,0.76
 10,1981,29,1981.79,0.79
 11,1981,30,1981.82,0.82
 12,1981,31,1981.85,0.85
 13,1981,32,1981.87,0.87
 14,1981,33,1981.90,0.90
 15,1981,34,1981.93,0.93
 16,1981,35,1981.96,0.96
 17,1981,36,1981.99,0.99
 18,1982,1,1982.01,0.01
 19,1982,2,1982.04,0.04
 20,1982,3,1982.07,0.07
 21,1982,4,1982.10,0.10
 22,1982,5,1982.13,0.13
 23,1982,6,1982.15,0.15
 24,1982,7,1982.18,0.18
 25,1982,8,1982.20,0.20
 26,1982,9,1982.23,0.23
 27,1982,10,1982.26,0.26

⋮

448,1993,35,1993.96,0.96
 449,1993,36,1993.99,0.99
 450,1994,1,1994.01,0.01
 451,1994,2,1994.04,0.04
 452,1994,3,1994.07,0.07

453,1994,4,1994.10,0.10
454,1994,5,1994.13,0.13
455,1994,6,1994.15,0.15
456,1994,7,1994.18,0.18
457,1994,8,1994.20,0.20
458,1994,9,1994.23,0.23
459,1994,10,1994.26,0.26
460,1994,11,1994.29,0.29
461,1994,12,1994.32,0.32
462,1994,13,1994.34,0.34
463,1994,14,1994.37,0.37
464,1994,15,1994.40,0.40
465,1994,16,1994.43,0.43
466,1994,17,1994.45,0.45
467,1994,18,1994.48,0.48
468,1994,19,1994.51,0.51
469,1994,20,1994.54,0.54
470,1994,21,1994.57,0.57
471,1994,22,1994.59,0.59
472,1994,23,1994.62,0.62
473,1994,24,1994.65,0.65
474,1994,25,1994.68,0.68

Acknowledgements

Thanks to Susan Campbell for valuable assistance in the reprojection of the PAL dataset.

References

- Gutman, G. and Ignatov, A., 1996, The relative merit of cloud/clear identification in the NOAA/NASA Pathfinder AVHRR Land 10-day composites. *Int. J. Remote Sensing*, **17**, 3295–3304.
- James, M.E. and Kalluri, S.N.V., 1994, The Pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring. *Int. J. Remote Sensing*, **15**, 3347–3363.
- Prince, S.D. and Goward, S.N., 1996, Evaluation of the NOAA/NASA Pathfinder AVHRR Land data set for global primary production modelling. *Int. J. Remote Sensing*, **17**, 217–221.
- Stowe, L.L., McClain, E.P., Carey, R., Pellegrino, P., Gutman, G.G., Davis, P., Long, C. and Hart, S., 1991, Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Advances in Space Research*, **3**, 51–54.
- Viovy, N. and Arino, O., 1992, The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *Int. J. Remote Sensing*, **13**, 1585–1590.