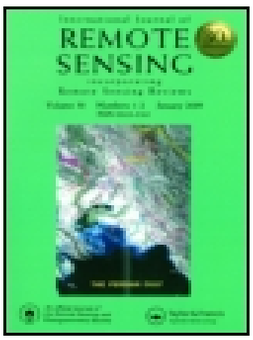

Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Noise estimation in NOAA AVHRR maximum-value composite NDVI images

L. R. EKLUNDH

To cite this article: L. R. EKLUNDH (1995) Noise estimation in NOAA AVHRR maximum-value composite NDVI images, Remote Sensing, 16:15, 2955-2962, DOI: 10.1080/01431169508954601

To link to this article: https://doi.org/10.1080/01431169508954601

Published online: 10 May 2007.

INT. J. REMOTE SENSING, 1995, VOL. 16, NO. 15, 2955–2962

Noise estimation in NOAA AVHRR maximum-value composite NDVI images

L. R. EKLUNDH

Department of Physical Geography, Lund University, Box 118, S-221 00 Lund, Sweden

Received 20 December 1994; in final form 24 April 1995)

Abstract. Noise has been estimated in Advanced Very High Resolution Radiometer (AVHRR) maximum-value composite Normalized Difference Vegetation Index (NDVI) products using a geostatistical method calculating the noise as the square root of the nugget variance in a semi-variogram. High noise values (20–50 per cent of the standard deviation) were found in individual scenes, whereas values close to zero were found in a nine-year average image. The higher levels of noise in individual images were present in humid areas particularly during the wet season, indicating that atmospheric effects and cloud contamination may be among the main contributing factors.

1. Introduction

The National Aeronautics and Space Administration (NASA) and the United Nations Food and Agriculture Organization (FAO) provide 10-day (decadal) composite images from the AVHRR (Advanced Very High Resolution Radiometer) of the NOAA (National Oceanic and Atmospheric Administration) series of polar orbiting satellites. These time series of images are used in early warning systems and food crop monitoring programmes such as the Famine Early Warning System of the United States Agency for International Development (USAID) and the FAO Global Information and Early Warning System (GIEWS). The data are used for identification of regions with imminent food shortages and to assess possible emergency food requirements (Hielkema 1992). Areas with suspected below-normal production are detected by comparing current images with images from previous years. It is of great interest to find out how much of the variation in these images can be attributed to real changes in the vegetation cover, and how much can be attributed to sensor noise, bi-directional effects, atmospheric influence, the resampling procedure and other factors not related to vegetation such as colour changes of the soil due to different moisture contents, etc.

This study aims at attempting to quantify noise in the imagery by analysing their spatial autocorrelation structures.

2. The NOAA NDVI data

The NOAA AVHRR system and data generated from it are described by Kidwell (1991). The Normalized Difference Vegetation Index (NDVI) is formed by combining the red and near-infrared bands of the data (Tucker *et al.* 1981). The index was early found to be related to green leaf biomass (Tucker 1979). NDVI data generated from NOAA AVHRR have been used from the early 1980s for assessing green

vegetation cover at regional and global scales (Prince and Justice 1991). The original Global Area Coverage (GAC) data have been composited temporally and spatially to create decadal maximum-value composite images (Holben 1986) with a spatial resolution of 7·6 km. A thermal cloud mask has been applied to the GAC data and the maximum NDVI value of the remaining part of the 10-day period has been kept to represent the period on a pixel-by-pixel basis. The images have been corrected for sensor drift and sensor differences by applying constants from Holben *et al.* (1990). In the current study 270 decadal NDVI composite images were used covering East Africa for the period 1982–1990. Other datasets suitable for global and regional vegetation monitoring also exist, such as the recent AVHRR Land Pathfinder data set. A review of this and other important AVHRR products is given in Townshend (1994).

3. Methods

Various methods exist for calculating noise in satellite data. A method by Gao (1993) is based on an evaluation of the local mean and standard deviation in small imaging blocks. The method relies on an assumption of additive noise only, and was not considered suitable for the current study. Aleksanina (1994) estimated noise in AVHRR IR channels by analysing spectrograms of quasi-homogeneous areas of images. The method used in this study was developed by Curran and Dungan (1989). It is called the geostatistical method since it uses a function commonly used in geostatistics, namely the semivariance. The semivariance can be used as a measure of the dependence of interpixel variability on distance and is defined as

$$\bar{S}^2 = (1/2m) \sum_{i=1}^{m} [z(x_i) - z(x_i + h)]^2 \quad (1)$$

where $\bar{S}^2$ is an unbiased estimate of the semivariance $\gamma(h)$ in the population, h is the lag distance between pairs of pizels, z is the signal of a single pixel x_i and m is the number of pairs of observations. As the distance between pixels increases so will the expected difference in pixel values, and thereby the semivariance. In a semivariogram the semivariance is plotted versus the lag. The value of $\gamma(h)$ as m approaches zero is called the nugget variance and is an estimate of the spatially independent variation in the image (Curran and Dungan 1989). This variation is normally considered to be made up mainly of uncorrelated noise (Jupp *et al.* 1989) but may also contain spatially uncorrelated scene information which is not noise (Curran and Dungan 1989). If the latter variation can be minimised the nugget variance will be close to zero in a noise-free image. As the distance increases the semivariance will increase until it reaches a maximum value, called the sill. In images with a global trend the semivariance will never reach the sill but continue to increase throughout the range of distances.

Curran and Dungan (1989) showed that the square root of the nugget variance can be used to estimate the standard deviation of the noise provided that the following assumptions are met:

1. Stationarity, i.e., no trend must be present in the data.
2. Isotropy, i.e., the semivariance must be uniform in all directions.
3. Fixed spatial resolution.
4. Scene does not contain random information.

5. Nugget variance is independent of the spatially dependent structural variance, i.e., the slope of the semivariogram does not influence the nugget variance.

The validity of these assumptions is discussed in § 5.

Data were analysed for five images. The first was an average 1982–1990 image created by averaging nine annual integrated NDVI images for the years 1982–1990. The other four were individual decadal images for the first decade in January, April, July and October 1990. The noise was calculated using the following method:

1. Four sub-scenes were selected from the original images. The aim was to find relatively homogeneous areas, yet sufficiently large for calculating the variograms. Statistics from the sub-scenes are presented in table 1.
2. 500 points were randomly distributed in each sub-scene and the semivariance calculated in a 13 by 13 pixel window around each point. For each window the centre pixel was compared with all other pixels, thus including all possible directions in the window. The 500 variograms were averaged to form one (one-dimensional) semivariogram showing the semivariance as a function of lag independent of direction, and one two-dimensional imager showing the directional dependence of the semivariance (the latter method was used by e.g., Stern 1985). The size of the pixel window required to establish the presence of a sill and any directional dependence may be different in other images; 13 by 13 pixels was considered appropriate in the present study.
3. Since it was noticed that the sampling pattern affected the resulting semivariogram step two was repeated 30 times to obtain a larger sample for the one-dimensional semivariograms. An alternative method to steps 2 and 3 would be to sample all pixels in the sub-scene.
4. The nugget variance at lag 0 was estimated by extrapolating linearly from a line fitted by least squares to the point clusters at lags 1 and 2. In studies where the characteristics of the semivariogram all the way to the sill are required, a line is normally fitted to the points at all lags. This was considered

Table 1. Descriptive statistics (NDVI) for the sub-scenes used in the analysis (μ = mean, σ = standard deviation).

	Somalia	Ethiopia	Uganda	Kenya
Climate zone	Arid to semi-arid	Humid highland	Humid	Semi-arid to humid
Rainfall (mm year^{-1})	100–500	1000–1500	1200–1600	500–1500
Number of cells	4800	3540	2508	6554
Mean 1982–90	μ: 0·12	μ: 0·35	μ: 0·35	μ: 0·26
	σ: 0·03	σ: 0·07	σ: 0·07	σ: 0·10
January 1990	μ: 0·13	μ: 0·33	μ: 0·35	μ: 0·45
	σ: 0·03	σ: 0·13	σ: 0·14	σ: 0·16
April 1990	μ: 0·13	μ: 0·27	μ: 0·43	μ: 0·39
	σ: 0·03	σ: 0·13	σ: 0·10	σ: 0·13
July 1990	μ: 0·08	μ: 0·42	μ: 0·39	μ: 0·29
	σ: 0·04	σ: 0·17	σ: 0·09	σ: 0·14
October 1990	μ: 0·11	μ: 0·56	μ: 0·41	μ: 0·22
	σ: 0·04	σ: 0·07	σ: 0·13	σ: 0·12

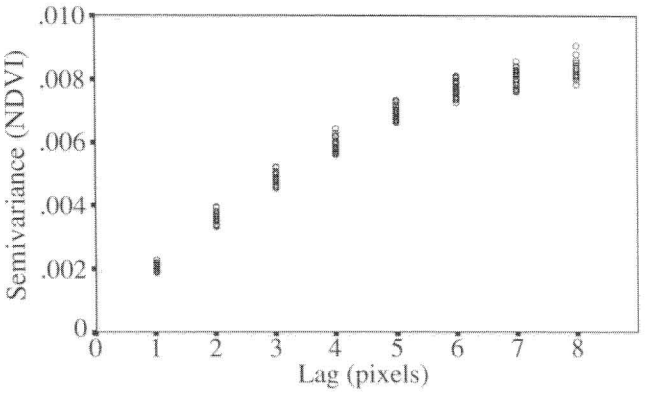

Figure 1. Semivariogram for Ethiopia, January 1990. Clusters at each lag contain 30 observations.

less appropriate for the current study since only the value of the nugget variance was of interest for the noise estimate.

5. The square root of the nugget variance was computed.

The signal-to-noise ratio (SNR) was calculated by dividing the mean value of the sub-scene by the square root of the nugget variance (Curran and Dungan 1989). These values were converted to dB by taking the logarithm and multiplying by 10. In order to estimate how much of the observed total variation in the images is due to noise the noise values were divided by the standard deviation of the corresponding sub-scenes and multiplied by 100.

4. Results

Figure 1 shows an example of one of the semivariograms produced from which the noise values were calculated. In figure 2 an example of one of the two-dimensional variograms is shown. The variogram was generated from the average

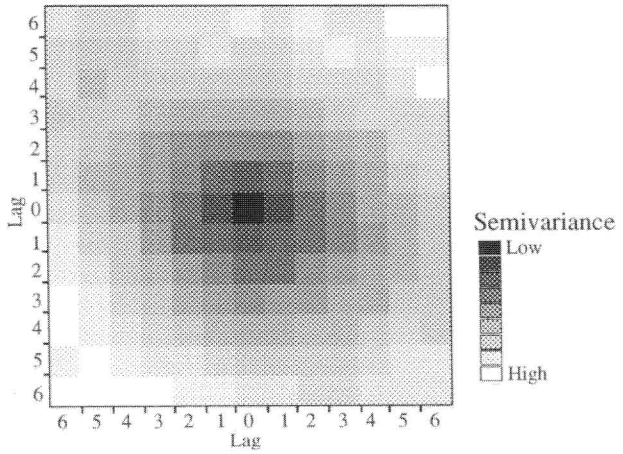

Figure 2. Two-dimensional-variogram for Ethiopia, average 1982–1990.

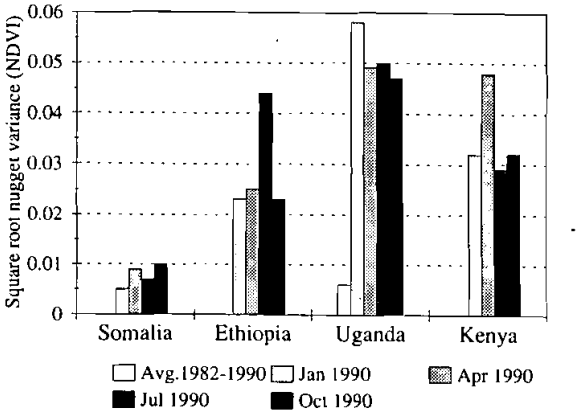

Figure 3. Noise calculated as the square root of the nugget variance for all sub-scenes.

1982–1990 image for Ethiopia and is an example of the anisotropic effect evident in some of the variograms. Anisotropy is also evident in the other mean 82-90 sub-scenes, but generally not for the individual January–October images 1990, except the Somalia July 1990 sub-scene which is strongly anisotropic. The resulting noise values for all images are presented graphically in figure 3. The SNR and noise as percentage of total sub-scene variation are presented in table 2. The noise values have been estimated to zero in three sub-scenes of the mean 1982-90 image (Somalia, Ethiopia and Kenya). Peak values in the 1990 scene occur for Somalia in October, Ethiopia in July, Uganda in January and Kenya in April. The lowest SNR values occur simultaneously with these (except for Kenya) but with small differences between the sub-scenes. The difference between the sub-scenes is larger when dividing the noise by the standard deviation. The noise values in the individual 1990 sub-scenes ranges from about 20 per cent of the standard deviation in the Somalia sub-scene to about 45 per cent in the Uganda sub-scene. The highest value recorded is for Uganda in July 1990 and is close to 54 per cent.

Table 2. Signal-to-noise ratio and noise expressed as percentage of the standard deviation for the four sub-scenes.

	SNR (dB)				Noise/standard deviation (%)			
	Somalia	Ethiopia	Uganda	Kenya	Somalia	Ethiopia	Uganda	Kenya
Average 1982–1990	—	—	17·6	—	0	0	8·3	0
January 1990	14·0	11·6	7·8	11·5	15·1	17·6	42·1	20·2
April 1990	11·5	10·3	9·4	9·1	27·1	19·8	47·4	36·5
July 1990	10·7	9·8	8·9	9·9	19·7	26·0	53·5	21·0
October 1990	10·3	13·9	9·4	8·3	24·5	32·7	35·8	26·1
Mean January–October 1990	11·6	11·4	10·6	9·7	21·6	24·0	44·7	26·0

5. Discussion

The results show that noise in the mean 82–90 sub-scenes is very low, probably due to the averaging of pixel values. In individual 1990 sub-scenes the values are much higher, exceeding 50 per cent in the highest case.

In order to assess the validity of the results it is important to study the necessary assumptions mentioned in §3. The first two of these are the assumptions of stationarity and isotropy. Whenever large regions with strong climate gradients are imaged, trends in the data are inevitable. Both the presence of trends and of anisotropy have been established in the data. One effect will be that the sill in the semivariogram is never reached. Following assumption five above, the nugget variance should be uncorrelated with the spatially dependent variance, and thus should not be affected by any trend. In practice, however, the nugget variance is estimated from the semivariogram and is likely to be inflated if a trend is present. In this study, the nugget variance was very close to zero in all the mean 1982–90 sub-scenes, despite the presence of both a trend and anisotropy. This indicates that the effect is very small.

The assumption of fixed spatial resolution generally does not hold for AVHRR data because of the large scan angles. Ground size of a pixel at the edges is about 12 times larger than at nadir (Moreno *et al.* 1992). In the NASA data compositing, the pixels at angles exceeding 43·3° have been excluded from the final product (concluded by Robinson 1991). This means that pixels are about three times larger at these angles than at nadir. The inclusion of these pixels decreases the semi-variance, and thereby the estimated noise, due to a spatial smoothing effect. It can be assumed that this influences individual images more than the long-term average since in the latter the influence of these angles has been averaged out for each pixel.

The assumption that all random information is noise cannot be met for certain when analysing sub-scenes of the size used in this study. Permanent random features, not changing from year to year, would increase the nugget variance in the average 1982–90 sub-scenes. Since nugget variance is very low in these images very few such features seem to be present. However, in the individual 1990 images seasonal changes in vegetation may lead to random variation affecting the nugget variance. Normalizing the noise estimate by the standard deviation of the sub-scene, as in table 2, would to some degree compensate for this.

The last assumption stating that nugget variance should be independent of the spatially dependent structural variance is by definition not true. However, because of the point-spread function of the sensor the effect of the violation will be minimal (Curran and Dungan 1989). In AVHRR data the point-spread effect is large due to significant overlapping between adjacent pixels along all the scan lines and particularly at large oblique viewing angles (Moreno *et al.* 1992).

The method used for extrapolating the semivariance to lag zero will also affect the result. The linear extrapolation used in this study led to small negative values at lag zero for three of the mean 82–90 image. Since this is not theoretically possible, the semivariances were set to zero. However, it highlights that problems exist with the method.

The peak noise values for Ethiopia and Kenya both occur simultaneously with the main rainfall season. In this season the number of missing data values due to clouds is far higher than during any other season. In Somalia the amount of missing data is negligible and in Uganda fairly high but without major differences between the seasons. There is strong reason to believe that large part of the noise in the sub-

scenes covering Ethiopia, Uganda and Kenya are due to atmospheric effects and cloud contamination. The noise values for Somalia are lower, and could probably be attributed mainly to other factors such as sensor noise, problems with the co-registration and sudden variations in surface colour due to burning.

Noise levels of individual images are high enough to question the usefulness of the images in studies where pixel by pixel comparisons are made. However, if data are averaged over a long time period very little noise will remain. The sub-scene over a semi-arid area (Somalia) showed significantly lower values than those of wetter areas, but was still high considering the intended use of the data. In the moist sub-scenes (Ethiopia, Uganda and Kenya) noise levels were very high, particularly during the wettest season. This indicates that the cloud filtering procedure used for these data needs to be improved. The amount of temporal (or possibly spatial) filtering needed to generate images with noise levels acceptable for monitoring purposes needs to be established in further studies.

Despite minor violations of the assumptions, the technique is relevant for AVHRR NDVI data. Users of maximum-value composite NDVI and similar datasets could apply the technique to their specific datasets for initial estimates of the noise contents and thereby usefulness of their data for monitoring purposes.

Acknowledgments

This study was made with financial support from the Swedish National Space Board. The NDVI data were obtained from USAID FEWS in co-operation with FAO Nairobi. I would like to thank Jonas Ardö for useful comments on the manuscript.

References

ALEKSANINA, M. G., 1994, Analysis of the noise characteristics of the AVHRR IR channels. *Soviet Journal of Remote Sensing*, **11**, 637–648.

CURRAN, P. J. and DUNGAN, J. L., 1989, Estimation of signal-to-noise: a new procedure applied to AVIRIS data. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **27**, 620–628.

GAO, B., 1993, An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers. *Remote Sensing of Environment*, **43**, 23–33.

HIELKEMA, J., 1992, Operational use of environmental satellite remote sensing and satellite communications technology for global food security and locust control by FAO. *World Space Congress, International Astronautical Federation (IAF)*, WSC Paper IAA-92-0397 (Washington D.C.: International Astronomical Federation), pp. 1–13.

HOLBEN, B. N., 1986, Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, **7**, 1417–1434.

HOLBEN, B. N., KAUFMAN, Y. J. and KENDALL, J. D., 1990, NOAA-11 AVHRR visible and near-IR inflight calibration. *International Journal of Remote Sensing*, **11**, 1511–1519.

JUPP, L. B., STRAHLER, A. H. and WOODCOCK, C. E., 1989, Autocorrelation and Regularization in Digital Images II. Simple Image Models. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **27**, 247–258.

KIDWELL, K. B., 1991, NOAA Polar Orbiter Data Users Guide, NOAA/NESDIS National Climate Data Center, Washington D.C., U.S.A.

MORENO, J. F., GANDÍA, S., and MELIÁ, J., 1992, Geometric integration of NOAA AVHRR and SPOT data: low resolution effective parameters from high resolution data. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, **30**, 1006–1014.

PRINCE, S. D., and JUSTICE, C. O., 1991, Editorial. *International Journal of Remote Sensing*, **12**, 1137–1146.

ROBINSON, T. P., 1991, Modelling the seasonal distribution of habitat suitability for armyworm population development in East Africa using GIS and remote sensing techniques. Ph.D. thesis. Department of Geography, University of Reading, U.K.

STERN, M., 1985, *Census from Heaven? Population Estimates with Remote Sensing Techniques.* Meddelanden från Lunds Universitets Geografiska Institution Avhandlingar XCIX. Ph.D. thesis, University of Lund, Sweden.

TOWNSHEND, J. R. G., 1994, Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction. *International Journal of Remote Sensing*, **15**, 3319–3332.

TUCKER, C. J., 1979, Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8**, 127–150.

TUCKER, C. J., HOLBEN, B. N., ELGIN, J. H., and MCMURTREY, J. E., 1981, Remote sensing of total dry matter in winter wheat. *Remote Sensing of Environment*, **11**, 171–189.