INT. J. REMOTE SENSING, 2000, VOL. 21, NO. 2, 401–407

Potential NDVI as a baseline for monitoring ecosystem functioning

D. M. STOMS

Institute for Computational Earth System Science, University of California, Santa Barbara, CA 93106, USA;
e-mail: stoms@geog.ucsb.edu

W. W. HARGROVE

Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

(*Received 17 February 1999; in final form 2 July 1999*)

Abstract. Baseline data are needed to determine the overall magnitude and direction of change in ecosystem functioning. This letter presents an approach to estimate potential NDVI from environmental variables and training data of actual NDVI in nature reserves. Patterns of deviations of actual NDVI from the baseline generally correspond with land-use types in the western United States.

1. Introduction

Numerous studies have shown that the Normalized Difference Vegetation Index (NDVI) derived from the Advanced Very High Resolution Radiometer (AVHRR) sensor data is related to ecosystem function, particularly net primary production (e.g. Goward *et al.* 1985, Box *et al.* 1989). NDVI has been used to monitor desertification (Tucker *et al.* 1991), land-use change (US Environmental Protection Agency 1997) and the effects of global warming in high latitudes (Myneni *et al.* 1997). Comparisons can only be made across the three decades for which satellite sensor data are available. Baseline data are not available to determine the total changes in NDVI and, hence, in ecosystem functioning. Our objectives in this letter are to implement the suggestion of Paruelo and Lauenroth (1995) about exploiting relationships with biophysical factors to model potential NDVI and then to examine the patterns of deviations of actual NDVI in terms of land uses.

2. Methods

The approach is to find training sites where actual NDVI approximates baseline values, formulate a model that best predicts these values, and apply that model to biophysical predictors to map potential NDVI. The study area covers three states (Washington, Oregon and California) in the western United States (figure 1). This regions spans 16 degrees of latitude and 10 degrees of longitude, with over 4000 m of topographic relief. Vegetation varies from cool, moist temperate rainforest to hot, arid deserts. Major land uses include cultivation, urbanization, logging and grazing. Thus the region provides a good sampling of environments and human land uses to

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online ©2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals/tf/01431161.html

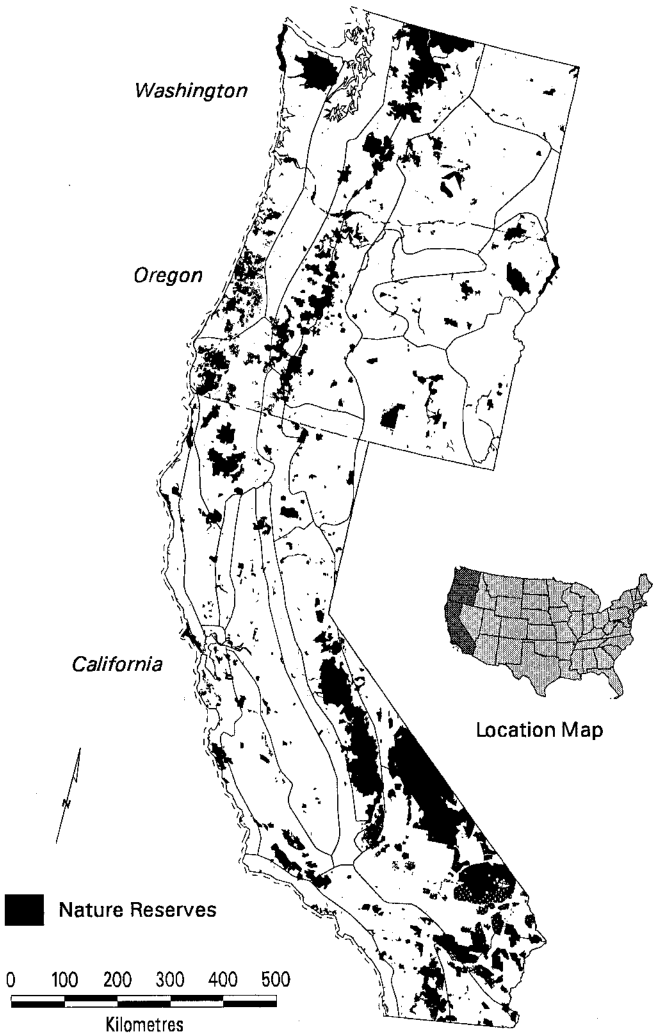

Figure 1. Managed areas compiled by the Gap Analysis Program and used as a source of training samples. Black lines delineate ecoregions.

explore the relationship between potential and actual NDVI and levels of environmental stress.

A set of predictor variables was compiled based on precipitation, temperature and soil characteristics at a resolution similar to the 1 km NDVI data (table 1) (Daly *et al.* 1994, Dodson and Marks 1997, Hargrove and Luxmoore 1998). The average of a series of nineteen 14-day NDVI composites (Eidenshink 1992) values for each 1 km pixel was computed as a measure of time-integrated NDVI (TI-NDVI) for 1990. The network of nature reserves has not experienced significant environmental stress from human activities (Paruelo and Lauenroth 1995). A random 5% sample of pixels was extracted from a map of these managed areas (figure 1) (Stoms *et al.* 1998) as training data for constructing the model.

Regression tree analysis (RTA, Venables and Ripley 1994) was used to build a predictive model of potential TI-NDVI. RTA was run ten times on random 90%

Table 1. Biophysical variables used in regression tree analysis of potential TI-NDVI.

Variable	Description
PPTANN	Mean annual precipitation, 1961–1990
PPTGROW	Mean precipitation in the growing season, 1961–1990
SOLGROW	Mean solar irradiance, 1983–1991
JANTMP	Mean January temperature, 1895–1993
JULTMP	Mean July temperature, 1895–1993
SEASTMP	JULTMP minus JANTMP, 1895–1993
DDHEAT	Degree-day heat sum, 1961–1991
DDCOOL	Degree-day cool sum, 1961–1991
EL	Digital elevation
AWC	Available soil water capacity (STATSGO)
OM	Total organic matter (STATSGO)
NITRO	Soil nitrogen (STATSGO)
WATDEP	Depth to seasonally-high water table (STATSGO)

samples and tested with the remaining 10%. The tree was pruned to the median number of terminal nodes in the ten runs. The RTA model was then applied to GIS layers of biophysical factors to produce a map of potential TI-NDVI for the study area. The actual TI-NDVI map was subtracted from the potential map, creating a difference image. A random 1% sample of pixels was analysed with respect to aggregated land-use classes (Loveland *et al.* 1991).

3. Results

Regression tree analysis of biophysical data accounted for 79% of the deviance in TI-NDVI for the training data (figure 2). The first division in the model was between areas of low and high PPTANN. Of the drier sites, those with slightly more precipitation were associated with higher TI-NDVI (Node B) in the high desert of eastern Oregon and Washington compared with the hot deserts of southern California (Node A). Of the wetter sites on the right side of the tree, colder JANTMP

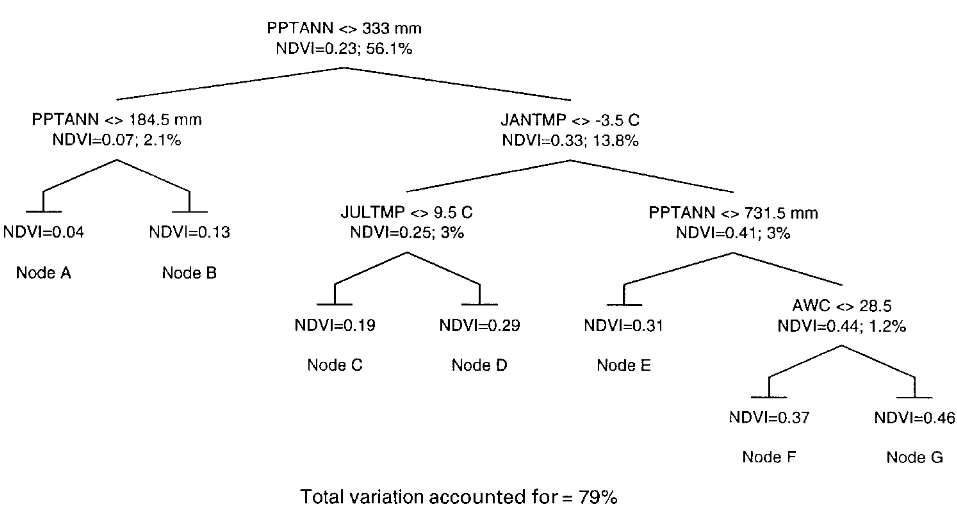

Figure 2. Regression tree of environmental predictors of potential TI-NDVI. Variable names are given in table 1.

corresponds to lower TI-NDVI in the interior mountains, particularly where JULTMP was also cool. For wet sites with mild winter temperatures, PPTANN divided the wetter north coast from the drier south coast. AWC was the final factor to split wetter sites. PPTANN alone accounted for more than 61% of the total variation in TI-NDVI. Temperature variables accounted for an additional 17%, and AWC added approximately 1%.

The difference image depicts the deviations of the actual TI-NDVI from that predicted by the RTA model (figure 3(*a*)). The negative deviations (greater TI-NDVI than predicted, in green) that immediately stand out are the large irrigated agricultural areas of central California and eastern Washington (figure 3(*b*)). Positive deviations (lower TI-NDVI than predicted, in red) are most visible in smaller patches, often urban areas.

Deviations varied somewhat as expected by land-use categories (figure 4). The Urban category tended to have positive deviations, indicative of a reduction in

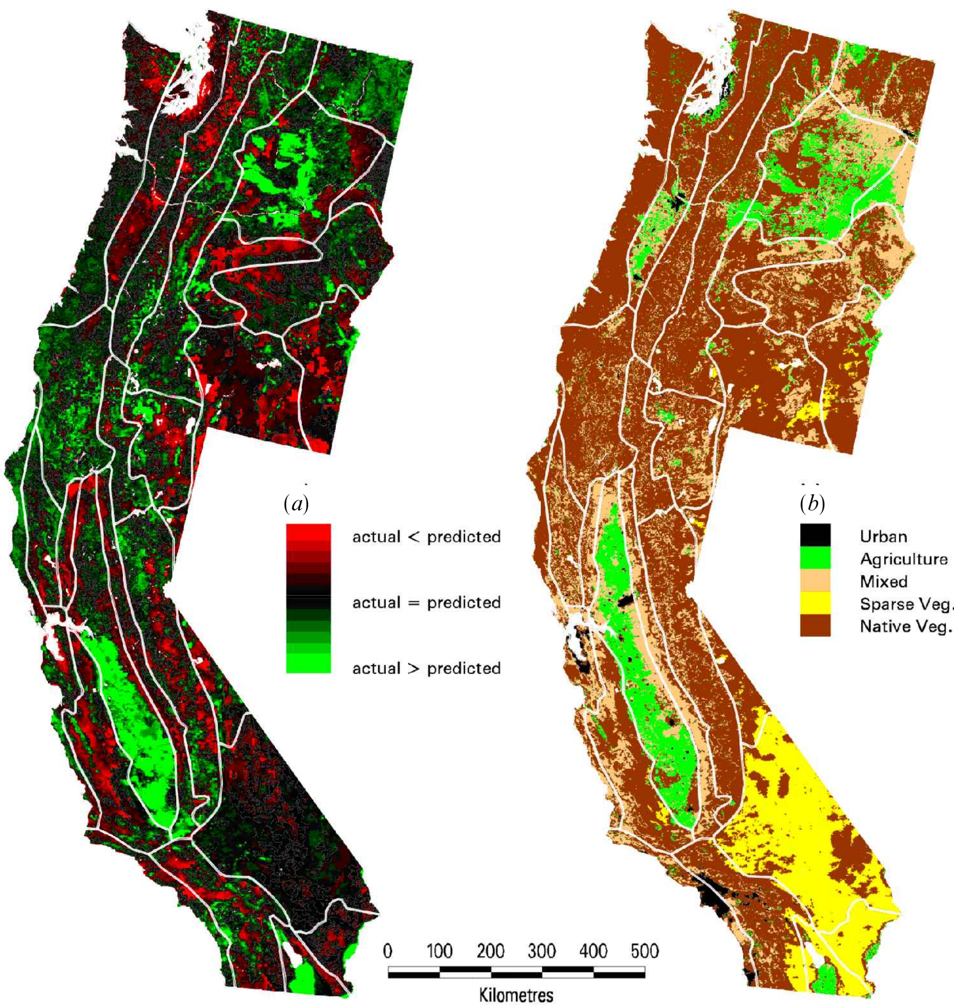

Figure 3. (*a*) Patterns of deviations between potential and actual time-integrated NDVI, and (*b*) land-use/cover type map.

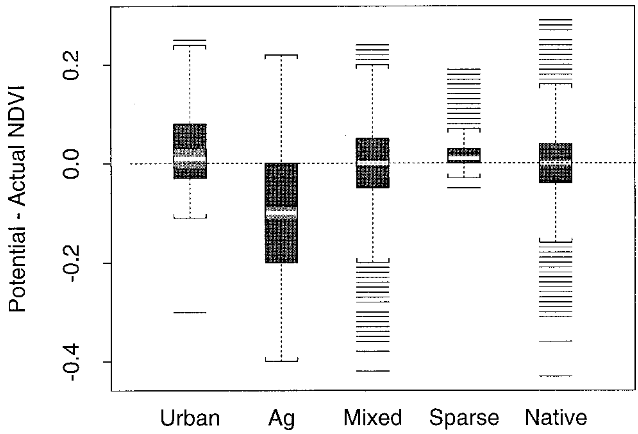

Figure 4. Boxplot of deviations between potential and actual time-integrated NDVI by land-use/cover type. The white bar indicates the median value of the deviation; the shaded bars mark the quartiles of the spread in values.

photosynthetic activity. Agricultural land had the widest distribution and most obviously negative deviation. Thus, agricultural lands in the study area tend to have greater TI-NDVI in the satellite imagery than would be expected from climatic and soils factors, owing to irrigation and fertilization. Mixed land, where agriculture and native vegetation are interspersed within pixels, looks much like native vegetation in the boxplot. Sparsely vegetated land shows a tendency towards slight positive deviations. This result may be because the model slightly overpredicts TI-NDVI in desert areas. The 'native' vegetation category includes all land-cover regions with little or no agricultural uses. The distribution of deviations for the native vegetation category was relatively narrow with a median near zero, as indicated by the position of the white bar.

4. Discussion and conclusions

Periodic satellite remote sensing facilitates monitoring ecosystem functioning. In this study, rather than using imagery to detect changes between two dates, we have developed a simple predictive model of potential TI-NDVI in the absence of human land-use effects. A model for the western United States used precipitation, temperature and available soil water capacity data, which captured most of the variation in TI-NDVI in undisturbed areas. Actual TI-NDVI deviated from potential TI-NDVI in the direction expected in response to urbanization and agriculture.

We chose to model potential TI-NDVI in this study because its ecological interpretation has been clearly established. Other vegetation indices have been developed to rectify NDVI's sensitivities or to provide complementary ecological information (e.g. Lambin and Ehrlich 1996, Huete *et al.* 1997, Bass *et al.* 1998) and should be considered for global monitoring. We believe that the type of empirical modelling outlined in this letter could also be used for similar indices.

While this approach appears to be promising for monitoring environmental stress, additional research is needed to corroborate these preliminary findings. Limitations of the present study can be categorized as those of AVHRR data, sampling, modelling and validation. The model was developed using NDVI composites from a single year. NDVI metrics can vary significantly between years in response to interannual differences in weather, local disturbance and sensor-related factors (Tucker *et al.* 1991, Myneni *et al.* 1997). Further research is needed to determine the optimal time period to represent baseline conditions. Training data for this study were extracted from areas being managed for natural ecosystem processes but their exact land uses and corresponding effects are not certain. Substantial effort will be required to identify enough sites that are truly free of human disturbance to provide an adequate sample of all environments. Finding sufficient numbers of sites will be problematic in some parts of the world. While capturing broad patterns in TI-NDVI, the model did not account for 21% of the deviance in the training data. Some of this may be due to local variation. Modelling may be more effective by individual ecoregions or by including ecoregions as categorical variables in the dataset. This pilot study only evaluated deviations between potential and actual TI-NDVI against mapped information on land use. Further work is needed to interpret the results with higher resolution map or field information to determine if apparent deviations truly reflect environmental stress. For all these reasons, we present the approach developed in this letter only as a promising initial step.

Acknowledgments

We thank Mary Cablk, Jeanie Sifneos, Denis White and Rusty Dodson for assistance. Although the research described in this article has been partially supported by the US Environmental Protection Agency through the Cooperative Agreement CR 823589-01-0 to Dr Frank Davis at the University of California, Santa Barbara, it has not been subjected to Agency review and therefore does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

References

Bass, B., Hansell, R., and Choi, J., 1998, Towards a simple indicator of biodiversity. *Environmental Monitoring and Assessment*, **49**, 337–347.

Box, E. O., Holben, B. N., and Kalb, V., 1989, Accuracy of the AVHRR vegetation index as a predictor of biomass, primary productivity and net CO_2 flux. *Vegetatio*, **80**, 71–89.

Daly, C., Neilson, R. P., and Phillips, D. L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**, 140–158.

Dodson, R., and Marks, D., 1997, Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Research*, **8**, 1–20.

Eidenshink, J. C., 1992, The 1990 conterminous U.S. AVHRR data set. *Photogrammetric Engineering and Remote Sensing*, **58**, 809–813.

Goward, S. N., Tucker, C. J., and Dye, D. G., 1985, North American vegetation patterns observed with the NOAA-7 Advanced Very High Resolution Radiometer. *Vegetatio*, **64**, 3–14.

Hargrove, W. W., and Luxmoore, R. J., 1998, A new high-resolution national map of vegetation ecoregions produced empirically using multivariate spatial clustering. *Proceedings of the 18th ESRI User Conference*, San Diego, California, 27–31 July 1998, available at < http://www.esri.com/library/userconf/proc98/PROCEED/TO350/-PAP333/P333.HTM> , last modified 1 April 1998.

Huete, A. R., Liu, H. Q., Batchily, K., and Van Leeuwen, W., 1997, A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, **59**, 440–451.

Lambin, E. F., and Ehrlich, D., 1996, The surface temperature-vegetation index space for land cover and land-cover change analysis. *International Journal of Remote Sensing*, **17**, 463–487.

Loveland, T. R., Merchant, J. W., Ohlen, D. O., and Brown, J. F., 1991, Development of a land cover characteristics database for the conterminous U.S. *Photogrammetric Engineering and Remote Sensing*, **57**, 1453–1463.

MYNENI, R. B., KEELING, C. D., TUCKER, C. J., ASRAR, G., and NEMANI, R. R., 1997, Increased plant growth in the northern high latitudes from 1981–1991. *Nature*, **386**, 698–702.

PARUELO, J. M., and LAUENROTH, W. K., 1995, Regional patterns of Normalized Difference Vegetation Index in North American shrublands and grasslands. *Ecology*, **76**, 1888–1898.

STOMS, D. M., DAVIS, F. W., DRIESE, K. L., CASSIDY, K. M., and MURRAY, M. P., 1998, Gap analysis of the vegetation of the Intermountain Semi-Desert Ecoregion. *Great Basin Naturalist*, **58**, 199–216.

TUCKER, C. J., DREGNE, H. E., and NEWCOMB, W. W., 1991, Expansion and contraction of the Sahara Desert from 1980 to 1990. *Science*, **253**, 299–301.

US ENVIRONMENTAL PROTECTION AGENCY, 1997, An Ecological Assessment of the United States Mid-Atlantic Region. EPA/600/R-97/130. Office of Research and Development, Washington, DC.

VENABLES, W. N., and RIPLEY, B. D., 1994, *Modern Applied Statistics with S-Plus* (New York: Springer-Verlag).