2005

A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the U.S. central plains

Tsegaue Tadesse

Jesslyn F. Brown

Michael Hayes

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ISPRS Journal of Photogrammetry & Remote Sensing 59 (2005) 244–253

PHOTOGRAMMETRY & REMOTE SENSING

www.elsevier.com/locate/isprsjprs

A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the U.S. central plains

Tsegaye Tadesse[a,*], Jesslyn F. Brown[b], Michael J. Hayes[a]

[a]*National Drought Mitigation Center, University of Nebraska-Lincoln, NE 68583-0749, USA*
[b]*SAIC, USGS/EROS Data Center, Sioux Falls, SD 57198, USA*

Received 1 March 2004; accepted 3 February 2005
Available online 29 March 2005

Abstract

Droughts are normal climate episodes, yet they are among the most expensive natural disasters in the world. Knowledge about the timing, severity, and pattern of droughts on the landscape can be incorporated into effective planning and decision-making. In this study, we present a data mining approach to modeling vegetation stress due to drought and mapping its spatial extent during the growing season. Rule-based regression tree models were generated that identify relationships between satellite-derived vegetation conditions, climatic drought indices, and biophysical data, including land-cover type, available soil water capacity, percent of irrigated farm land, and ecological type. The data mining method builds numerical rule-based models that find relationships among the input variables. Because the models can be applied iteratively with input data from previous time periods, the method enables to provide predictions of vegetation conditions farther into the growing season based on earlier conditions. Visualizing the model outputs as mapped information (called VegPredict) provides a means to evaluate the model. We present prototype maps for the 2002 drought year for Nebraska and South Dakota and discuss potential uses for these maps.

Keywords: Drought; Data mining; Vegetation stress; Satellite metrics; Prediction; NDVI

1. Introduction

Drought is a natural hazard that impacts economic, social, and environmental aspects of society. In the agricultural sector, it is one of the dominant causes of crop loss in USA (Wilhite, 2002). In many cases, agricultural losses are increased by reductions in livestock production and disruptions in the food supply chain (Goddard et al., 2003). In recent years, droughts caused billions of dollars in damages/losses in many states. For example, in 2002, the estimated agricultural losses exceeded US$1 billion in each of

* Corresponding author. Tel.: +1 402 472 3383; fax: +1 402 472 6614.
E-mail address: ttadesse2@unl.edu (T. Tadesse).

0924-2716/$ - see front matter
doi:10.1016/j.isprsjprs.2005.02.003

the states of Nebraska, Colorado, Kansas, and South Dakota (Hayes et al., 2004). The United States Federal Emergency Management Agency (FEMA) has estimated that drought events are responsible for annual economic losses of US$6–8 billion (FEMA, 1995).

Given the seriousness of these losses and the severity of recent droughts across the country, policy-makers in the USA have significant interest in monitoring and predicting drought events. To monitor drought, decision-makers at the administrative and grass-roots levels need timely and accurate information about the spatial and temporal dimensions of droughts. This information helps officials and farmers to be more proactive in managing drought risk (Wilhite, 2002). Furthermore, drought impacts can be reduced through better understanding of drought and identifying the appropriate drought indicators for an early warning system. This includes providing decision-makers with timely drought products (e.g., maps and data) that identify the frequency, severity, and spatial extent of drought.

In the past, climate and meteorological data have been the primary sources for drought information used to support decision-making. However, more recently satellite observations have proved to be a valuable source of timely, spatially continuous data with improved detail for monitoring vegetation dynamics over large areas. Many prior studies of vegetation conditions base analyses on numerical transforms known as vegetation indices (VI). These indices have been used for studying vegetation characteristics over large areas since the 1970s (Rouse et al., 1974; Tucker, 1979). The advantages of using VIs rather than strictly spectral observations include minimizing soil and other background effects, reducing data dimensionality, providing a degree of standardization for comparison, and enhancing the vegetation signal (Curran, 1981; Goward, 1989; Malingreau, 1989). One of the more commonly used VIs, the Normalized Difference Vegetation Index (NDVI), takes advantage of the reflective and absorptive characteristics of plants in the red and near-infrared portions of the electromagnetic spectrum.

Various studies have demonstrated the utility of satellite measurements for observing and monitoring drought and provide analyses of the relationships between climate variables (e.g., precipitation) and satellite-derived VIs (Di et al., 1994; Yang et al., 1998; Ji and Peters, 2003). McVicar and Bierwirth (2001) investigated the utility of satellite data as a drought assessment tool for the 1997 drought in Papua New Guinea. They found a strong correlation (r^2=0.809) between accumulated rainfall and an integrated measurement of surface temperature (T_s) and NDVI over meteorological stations. In another study, Yang et al. (1998) investigated the relationships among several climate parameters (including growing season precipitation) and an annual integration (or summation) of NDVI over grasslands in the U.S. Great Plains. When examined over all grasslands in the analysis, model results showed a significant positive correlation between the time-integrated NDVI and spring and summer precipitation. Ji and Peters (2003) showed significant correlations between monthly NDVI and the SPI during the growing season over four states in the U.S. central plains. Even though this study was based on spatially averaged NDVI and SPI data (calculated over climate divisions), they found NDVI to be an affective indicator of moisture and vegetation condition.

Additional studies have presented analyses of droughts in the USA, Africa, South America, and Asia illustrate how derivatives of the NDVI can improve the ability to observe drought in time-series satellite data (Kogan, 1995; Liu and Kogan, 1996; Unganai and Kogan, 1998). The Vegetation Condition Index (VCI), a ratio of NDVI collected in a given period compared to its historical range (maximum minus minimum) for the same period during multiple years of record, is used to map drought patterns (Kogan, 1995). Peters et al. (2002) demonstrated the potential of a measure called the Standardized Vegetation Index for drought monitoring over the U.S. Great Plains and presented six monthly maps for the year 2000.

In this paper, we introduce a prototype vegetation stress map called *VegPredict* that depicts vegetation conditions several weeks in advance. This map product is created using data mining techniques that integrate satellite, climate, and other environmental data sets. The data mining technique maximizes the information contained in traditional drought indicators and integrates it with satellite-based greenness measures from the Advanced Very High Resolution Radiometer (AVHRR) processed at the USGS EROS Data Center (Eidenshink, 1992). This paper demonstrates the potential use of data mining for drought research and presents map outputs over the central states of Nebraska and South Dakota for the 2002 drought year.

2. Using data mining for drought mapping and prediction

Data mining is a technique that uses a variety of data analysis tools to discover patterns and relationships of physical variables. This technique has shown promise for analysis and prediction in multiple disciplines bringing together techniques from machine learning, pattern recognition, statistics, databases, and visualization to address the issue of information extraction from large databases (Cabena et al., 1998; Groth, 1998).

Studies in ecological research have also introduced data mining techniques and found that it is a powerful tool in addressing complex ecological problems handling both numeric and categorical data (De'ath and Fabricius, 2000). Recent studies have shown that, although drought effects on vegetation result from complex atmospheric and biophysical phenomena, data mining provides mechanisms for understanding drought characteristics in space and time (Tadesse et al., 2004; Harms et al., 2002). These studies illustrate the potential of data mining for drought analysis and prediction.

2.1. Rule-based predictive regression-tree model

In this study, *Cubist*[1] data mining software was used to generate models from a combination of satellite, climate, and biophysical data. The technique is generally referred to as regression-tree modeling. *Cubist* analyzes data and generates rule-based linear models that are a collection of rules, each of which is associated with a linear expression for computing a target value. The user determines the dependent and independent variables. To get a more reliable estimate of accuracy, the data are automatically divided into a number of folds to validate the rules. In this study, the data are divided into five blocks of almost equal size and target value distribution. For each block, *Cubist* constructs a model from the cases in the remaining blocks and tests on the cases in the hold-out block. In this way, each case is used once as a test case. The accuracy of a model produced from all the cases is estimated by averaging results on the hold-out cases (Rulequest Research, 2003). The final output includes a summary of the average error of the prediction and correlation coefficient values. The correlation coefficient (r) measures the statistical agreement between the cases' actual values of the target attribute and those values predicted by the model.

2.2. Satellite, climate, and biophysical data

To predict future vegetation condition, we used data from a satellite source, climate-based drought indices, and biophysical variables For the climate data, drought indicators were calculated for weather station locations. These became the model generation locations. The satellite data, and many of the biophysical variables, were extracted using GIS techniques for the same weather station locations.

2.2.1. Climatic indices

The Standardized Precipitation Index (SPI) and the Palmer Drought Severity Index (PDSI) were selected to define and quantify precipitation deficits (McKee et al., 1994; Palmer, 1965). The SPI and PDSI were calculated at 14-day intervals to match the temporal resolution of the satellite data. For this study, we used the Self-Calibrated PDSI that provided improvement as it accounts for climate and soil characteristics of each weather station that was not the part of the original PDSI algorithm (Wells et al., 2004).

2.2.2. Satellite data

Satellite-derived measure of vegetation stress, the Percent Average Seasonal Greenness (PASG) was calculated based on smoothed temporal NDVI curve characteristics. For this project, our source is AVHRR NDVI calculated on a 14-day time step. The start of the season (SOS) and end of season are used in the calculations for seasonal greenness (SG). To identify the SOS, we discover a well-defined trend change in the NDVI vector using a delayed or backward-looking moving average (Reed et al., 1994). This method

[1] Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the USA Government.

identifies a well-defined trend change in the NDVI (and vegetation vigor). The SG is calculated as a daily integration of the NDVI (interpolated from 14-day NDVI) above the NDVI baseline between the median SOS day and all other periods in the growing season. For vegetation monitoring, the SG at each pixel for any given period is compared to the mean of the same time period from the historical (from 1989 to 2002) database. The measure is expressed as a percentage by the formula: PASG=(current SG/mean SG)*100.

2.2.3. Biophysical variables

A set of additional biophysical variables was chosen based on assumptions underlying vegetation behavior. For all variables listed in Table 1, the dominant (or mean) value was calculated for a 9-km^2 window surrounding each weather station. Land-cover type is used in our model to identify the influence of the vegetation cover and its impact on drought. The soil available water capacity indicates how well the soil can hold moisture based on soil type. The percent of land in farms in irrigation indicates sensitivity to drought. The ecoregions designation provides the responsiveness of ecological resources to drought conditions. These variables were chosen because they influence vegetation response to climate condition; however, the list is not exhaustive.

Finally, the above described climate, satellite, and biophysical data variables were combined to generate rules that show the relationships among the variables. After generating regression tree models from data generated at weather station locations, we converted the models into map output. The rule-based models estimate a value of the dependent variable for each station (or pixel) depending on the independent variables. After the rules are generated, the map is produced based on the value for each station (pixel). The rule for each station (pixel) is selected if the case satisfies all the conditions of the rule for that specific station. Consider the following two rules to illustrate the process to determine the value for each pixel.

Rule 1: If PDSI<−1.5, land-cover is grassland and percent irrigation (Pirrg) is less than 10%, then VegPredict=46.51+3.9spi−1.14pdsi−0.077Pirrg; and
Rule 2: If PDSI≤−2.5, land-cover is pasture, available water capacity (awc) of soil is 68, then VegPredict=20.5875−0.038awc+0.031Pirrg+0.13 pdsi.

The coefficient (weight) for each variable in the regression tree model rule is determined by the regression tree algorithm based on the historical data. Moreover, all of the variables at a given time and place are not used in a single rule. The number of rules that are generated can be automatically determined by the data mining algorithm or can also be determined by the user. For example, 40 rules may be generated for the predictive models. For each pixel, the decision rules are chosen automatically when the conditions are satisfied in calculating the predictive values. Fig. 1 shows the flow chart used to model and produce the output maps. These output maps (VegPredict maps) are produced using research code developed at the EROS Data Center. This code converts the *Cubist*-derived rules and applies them to each pixel in binary geospatial imagery. The following section demonstrates the process, modeling and production of the time-dependent VegPredict maps as a case study in the central plains for the 2002 drought year.

3. Modeling and predicting drought related vegetation stress for the central plains: a case study of the 2002 drought

Data mining can be used for prediction using the time lag relationships of the variables. In this case

Table 1
Biophysical data sets

Data	Source	Date	Reference
Land-cover	National Land Cover Database	1990–1992	Vogelmann et al. (2000)
Soil available water capacity	State Soil Geographic Data Base	NA	USDA (1994)
Percent of land in farms in irrigation	Census of Agriculture	1997	USDA (1997)
Ecoregions	Environmental Protection Agency	1987 (revision, 3/2000)	Omernik (1987)

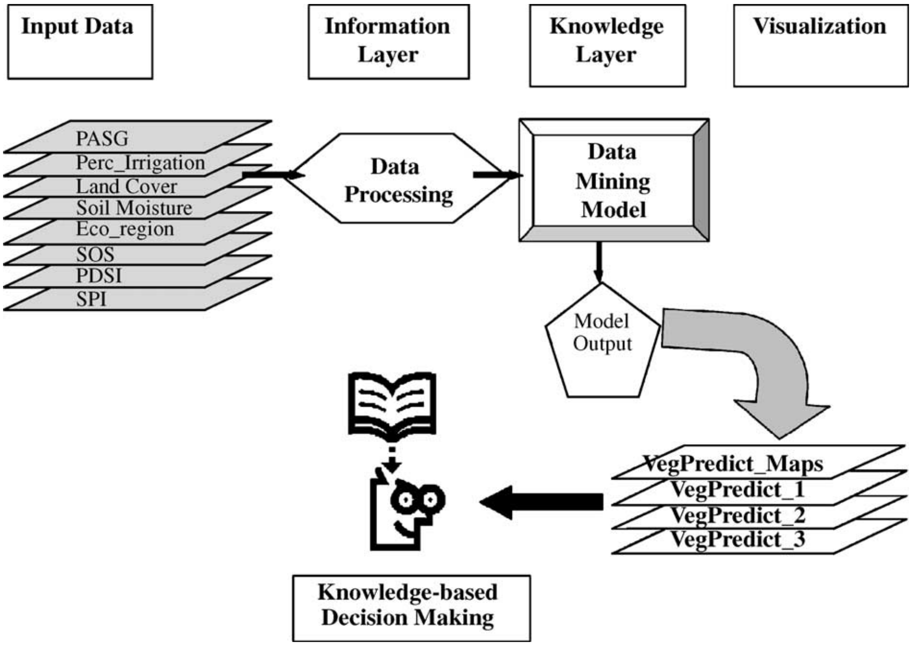

Fig. 1. Flow-chart of the process used to produce the VegPredict maps for decision-making.

study, we predict drought impact on vegetation stress based on the values of the PASG using climate and other biophysical data as independent variables. The PASG is selected as the dependent variable mainly because vegetation stress typically occurs after a precipitation (water supply) deficit affects plant growth. The predictive regression-tree models were developed based on historical data from 224 weather stations in South Dakota and Nebraska for 14 years (i.e., from 1989 to 2002).

Regression tree models were built for each of three phases described in Table 2. First rule-based models were generated to identify relationships between the vegetation condition during the growing season and the other climatic and biophysical variables using no (zero) time lag. In this step, we predict current vegetation condition without the PASG as an input variable. Next, we generated rules that showed relationships between climate data and satellite vegetation data (i.e., PASG) observed after 2, 4, and 6 weeks to investigate the prediction capabilities. These rule-based models serve as predictive mechanisms for the respective time lags (e.g., 2, 4, and 6 weeks).

After generating the predictive rules, we applied the models to the 2002 drought year to demonstrate its application in assessing and predicting the seasonal vegetation stress for Nebraska and South Dakota. The inputs for generating these maps included SPI, PDSI, and PASG data in raster formats for 12 periods in the growing season defined from April 19th to October 3rd, 2002. For each period, geospatial raster maps that show the predicted severity and spatial extent of vegetation stress were generated. Fig. 2(a) shows the VegPredict map generated for the period ending August 22nd, 2002. This period was selected as an example to demonstrate the impact of drought in the middle of the growing season.

Table 3 summarizes the results of the iterative process of the model runs for different time lags (i.e., prediction periods), which include corresponding correlation (r^2) and average error values. Model results showed strong relationships between the

Table 2
Three phases of the growing season

Phase	Phenological stage	Period
Phase I	Greenup, early growth	Spring (April 19 to June 13)
Phase II	Maturity, peak growth	Summer (June 14 to August 22)
Phase III	Senescence, harvest	Fall (August 23 to October 3)

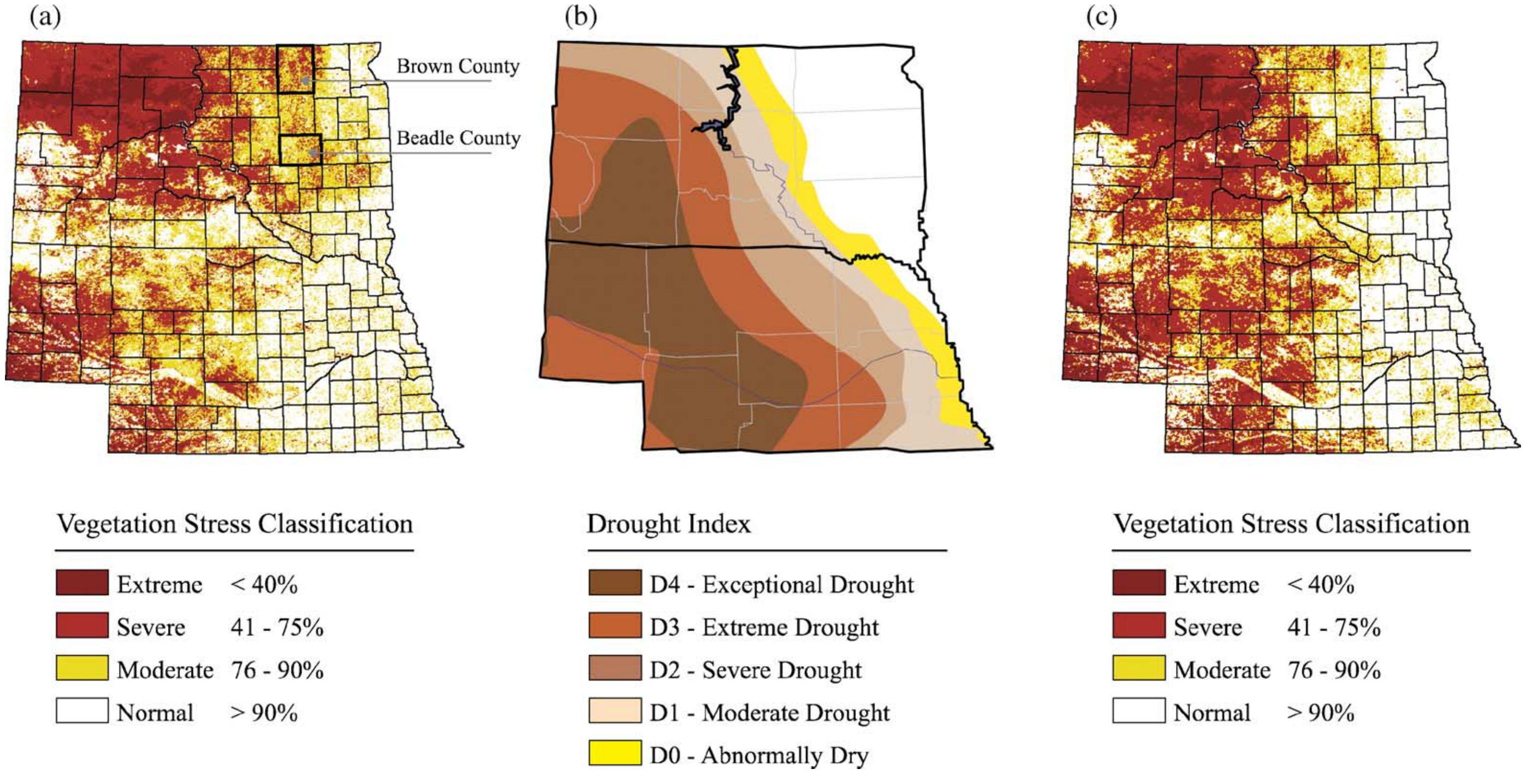

Fig. 2. (a) The VegPredict map that was predicted 6 weeks earlier for the biweekly period ending 22nd August 2002; (b) drought monitor map for 27 August 2002. This map is for Nebraska and South Dakota and was extracted from the U.S. Drought Monitor (NDMC, USDA, and NOAA, http://www.drought.unl.edu/dm/); (c) the Percent Average Seasonal Greenness (PASG) derived from the actual satellite data for the biweekly period ending 22nd August 2002.

Table 3
Errors (unitless) and correlation coefficients of rule-based models for VegPredict for different time-lag periods

Time lags	For period	Average |error|	r^2
Zero lag	19th April to 13th June	15.53	0.44
	14th June to 22nd August	5.75	0.81
	23rd August to 3rd October	5.07	0.72
Two-week prediction	3rd May to 27th June	8.58	0.67
	28th June to 5th September	1.99	0.96
Four-week prediction	16th May to 11th July	8.61	0.66
	12th July to 19th September	3.48	0.85
Six-week prediction	31st May to 25th July	8.14	0.66
	26th July to 3rd October	3.91	0.85

vegetation condition and the climatic and biophysical parameters.

The relationship of the climate data with the vegetation stress showed very low values of r^2 (0.44) for the early phenological phase. This may be due to instability in the PASG close to the start of the growing season. However, the relationship is stronger during the periods of maturity and senescence (period prior to harvest and leaf-drop) since vegetation activity is more stable and predictable during these periods if the climatic conditions are known. This was evident by the r^2 values of 0.81 and 0.72 for the summer and fall periods, respectively. The authors duly note that these model results bear the influence of spatial and temporal autocorrelation. For example, stations with close distance to each other may increase the r^2 values because of their proximity observing similar climatic data. Efforts to account for these effects are in progress.

We have examined various techniques for the validation and evaluation of the VegPredict maps. One of the methods used is comparing the geospatial output with the U.S. Drought Monitor (USDM), currently used by many Federal, State, and local agencies to monitor drought (Svoboda et al., 2002). Other methods include comparing VegPredict with PASG derived from actual satellite data for the same periods, and validating the extent of the vegetation stress with the crop yield data.

3.1. Comparing VegPredict with the USDM and the actual satellite-derived PASG

In a comparison between the VegPredict map for period ending August 22, 2002 and the USDM for the same period (Fig. 2(b)), we see both significant similarities as well as differences in the spatial drought patterns. Because of relatively higher spatial resolution (1 km^2), more detailed patterns of vegetation stress are identified in the VegPredict compared to the USDM. The main difference between VegPredict and the USDM for this period is that the USDM delineated eastern South Dakota as an area showing no drought. We do not have a definitive answer for this difference; however, it may be due to the high intensity rain events that occurred in late July and early August in this area influenced the USDM depiction to miss the vegetation stress. The other discrepancies between the USDM and the VegPredict map are in most areas of eastern Nebraska. These differences may result from a considerable amount of irrigated agriculture that influences the decision rules generated for these areas.

Fig. 2(a) and (c) show many similarities between the VegPredict and the actual satellite-derived PASG. However, we observed some differences in the intensity of the vegetation stress in the eastern parts of Nebraska and South Dakota. One possible explanation for these differences could be that the heavy thunderstorms, which brought drought-easing rain for these areas, were not anticipated in the predictive rules and data in early August 2002. Generally, the strong similarity of the two figures indicates how well the VegPredict maps can predict vegetation stress. This may be helpful to the agricultural sector, especially in assessing and predicting pasture and rangeland conditions during the growing season.

3.2. Comparison with detrended crop yields for selected counties

One source for evaluation of VegPredict involves using the National Agricultural Statistics Service (NASS) county-level crop yield data for conformation of drought conditions (NASS, 2004). Detrending and normalization of crop yield data is necessary to minimize the effects of technological advancement on increasing production through the years. After detrending and normalizing, we can make more realistic comparisons with crop yield data through time. In this study, we selected Brown and Beadle counties in eastern South Dakota where we observed differences in drought patterns between the USDM and VegPredict. The USDM depicted these counties in eastern South Dakota as areas having no drought impacts (Fig. 2(b)). In contrast, VegPredict indicated that these counties were undergoing vegetation stress because of drought (Fig. 2(a)).

Based on the 1997 Census of Agriculture data, the crops that have the largest acreage in Brown and Beadle counties are corn and wheat, respectively (NASS, 2004). Fig. 3(a) and (b) show the detrended and normalized corn and wheat production of Brown and Beadle counties. Both figures show 2002 as the lowest crop yields in both counties. This was confirmed by patterns in the VegPredict map but not in the USDM. The actual satellite data (PASG) for the same period confirmed vegetation stress within these counties (Fig. 2(c)).

4. Future challenges and directions

At this point, we are faced with the challenge of objectively validating model output across the study

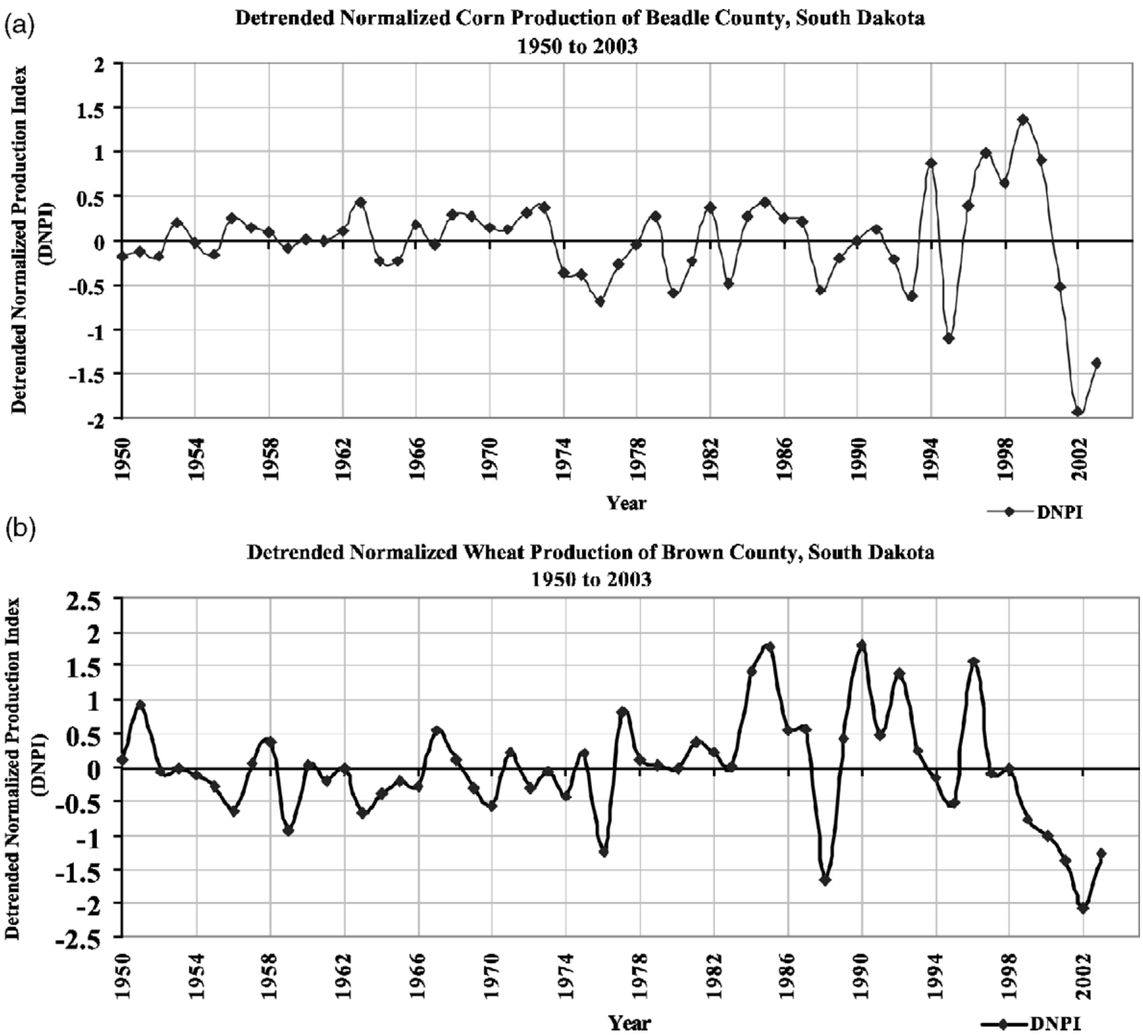

Fig. 3. (a) Detrended normalized corn production of Beadle County, South Dakota; (b) detrended normalized corn production of Brown County, South Dakota.

area. This task is challenging because VegPredict delivers continuous spatial coverage and is inherently finer in spatial detail than other commonly available drought indicators. Thus, we are pursuing data evaluation through several avenues.

The first method is comparing the drought patterns within agricultural areas utilizing county-level agricultural production statistics for additional counties instead of only selected counties, which were completed for the purpose of this study. The second approach is focused on determining the best way to deliver drought information to the public. This approach centered on holding several citizen panels where potential users will be asked to provide input on VegPredict and a web-enabled map application tool. Through citizen panels, we will assemble recommendations from the participants, which may include farmers, ranchers, extension agents, and experts from land management agencies. This information will then become a component in future development for both web-mapping tools and the content and effectiveness of the near-real time drought information.

5. Conclusions

In this study, we have introduced a new drought prediction product that (a) has finer spatial detail than previously available, (b) provides current drought status and predicts future drought (vegetation) conditions, and (c) utilizes data mining and image processing techniques. The techniques appear to have wide applicability, and have potential for other areas, both in the USA and internationally. At this stage, evaluation and validation of model products are the highest priority. Because most existing drought information is much coarser in spatial and temporal detail, appropriate for only a subset of land-cover types (e.g., county-based agricultural statistics), and usually surrogate, validating these results remains a challenge.

The ultimate value of this kind of drought vegetation stress prediction is yet to be realized; however, the improved spatial resolution of this map product has the potential to provide detailed information up to 6 weeks earlier, providing information for decisions made at county and even community level.

Acknowledgements

This research was performed in part by the Raytheon and SAIC Corporations under U.S. Geological Survey Contracts 1434-CR-97-40274 and 03CRCN0001 and by the National Drought Mitigation Center under U.S. Geological Survey Grant 02CRGR0003.

References

Cabena, P.H., Stadler, R., Verhees, J., Zanasi, A., 1998. Discovering Data Mining: From Concept to Implementation. IBM, New Jersey. 195 pp.

Curran, P.J., 1981. Multispectral remote sensing for estimating biomass and productivity. In: Smith, H. (Ed.), Plants and the Daylight Spectrum. Academic Press, London, pp. 65–99.

De'ath, G., Fabricius, K.E., 2000. Classification and regression trees—a powerful yet simple technique for ecological data analysis. Ecology 8 (11), 3178–3192.

Di, L., Rundquist, D.C., Han, L., 1994. Modeling relationships between NDVI and precipitation during vegetative growth cycle. International Journal of Remote Sensing 15 (10), 2121–2136.

Eidenshink, J.C., 1992. The 1990 conterminous U.S. AVHRR data set. Photogrammetric Engineering and Remote Sensing 58 (6), 809–813.

Federal Emergency Management Agency (FEMA), 1995. National mitigation strategy—partnerships for building safer communities. FEMA, Washington, D.C. 26 pp.

Goddard, S., Harms, S.K., Reichenbach, S.E., Tadesse, T., Waltman, W.J., 2003. Geospatial decision support for drought risk management. Communication of the ACM 46 (1), 35–37.

Goward, S.N., 1989. Satellite bioclimatology. Journal of Climate 2 (7), 710–720.

Groth, R., 1998. Data Mining: A Hands-on Approach for Business Professionals. Prentice Hall Inc., New Jersey. 264 pp.

Harms, S.K., Deogun, J., Tadesse, T., 2002. Discovering sequential association rules with constraints and time lags in multiple sequences. In: Hacid, M.S., Ras, Z.W., Zighed, D.A., Kodratoff, Y. (Eds.), Proc. the 13th International Symposium on Methodologies for Intelligent Systems, Lyon, France, June 27–29. Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence, vol. 2366. Springer, Berlin, pp. 432–441.

Hayes, M.J., Svoboda, M.D., Knutson, C.L., Wilhite, D.A., 2004. Estimating the economic impacts of drought. 84th AMS Annual Meeting, Proc. the 14th Conference on Applied Climatology, Seattle, Washington, USA, 10–16 January (on CD-ROM).

Ji, L., Peters, A.J., 2003. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. Remote Sensing of Environment 87 (1), 85–98.

Kogan, F.N., 1995. Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. Bulletin of the American Meteorological Society 76 (5), 655–668.

Liu, W.T., Kogan, F.N., 1996. Monitoring regional drought using the vegetation condition index. International Journal of Remote Sensing 17 (14), 2761–2782.

Malingreau, J.P., 1989. The vegetation index and the study of vegetation dynamics. In: Toselli, F. (Ed.), Application of Remote Sensing to Agrometeorology. ECSC, Brussels, pp. 285–303.

McKee, T.B., Doesken, N.J., Kleist, J., 1994. Drought monitoring with multiple time scales. Preprints, 9th conf. on Applied Climatology. Amer. Meteor. Soc., Dallas, TX, pp. 233–236.

McVicar, T.R., Bierwirth, P.N., 2001. Rapidly assessing the 1997 drought in Papua New Guinea using composite AVHRR imagery. International Journal of Remote Sensing 22 (11), 2109–2128.

NASS (National Agricultural Statistics Service), 2004. Historical crop yield data. http://www.usda.gov/nass/pubs/histdata.htm (accessed 3 February 2005).

Omernik, J.M., 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77 (1), 118–125.

Palmer, W.C., 1965. Meteorological drought. U.S. Department of Commerce, Weather Bureau, Research Paper, vol. 45. 58 pp.

Peters, A.J., Walter-Shea, E.A., Lei, J., Vina, A., Hayes, M., Svoboda, M.R., 2002. Drought monitoring with NDVI-based standardized vegetation index. Photogrammetric Engineering and Remote Sensing 68 (1), 71–75.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science 5 (5), 703–714.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. NASA/GSFC Type II Final Report, Greenbelt, Maryland, 371 p.

Rulequest Research, 2003. An overview of Cubist. http://www.rulequest.com/cubist-win.html (accessed on 3 February 2005).

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Thinker, R., Palecki, M., Stooksbury, D., Miskus, D., Stephens, S., 2002. The drought monitor. Bulletin of the American Meteorological Society 83 (8), 1181–1190.

Tadesse, T., Wilhite, D.A., Harms, S.K., Hayes, M.J., Goddard, S., 2004. Drought monitoring using data mining techniques: a case study for Nebraska, USA. Natural Hazards Journal 33 (1), 137–159.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8 (2), 127–150.

Unganai, L.S., Kogan, F.N., 1998. Drought monitoring and corn yield estimation in southern Africa from AVHRR data. Remote Sensing of the Environment 63 (3), 219–232.

USDA (U.S. Department of Agriculture), 1994. State Soil Geographic (STATSGO) Data Base: data use information. Natural Resource Conservation Service, Soil Conserv. Serv., Miscellaneous Publication Number 1492.

USDA (U.S. Department of Agriculture), 1997. Census of Agriculture, Geographic Area Series. National Agricultural Statistics Service, Washington D.C. (on CD-ROM). Available also at http://www.nass.usda.gov/census/index1997.htm (accessed 3 February 2005).

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, N., 2000. Completion of the 1990s national land cover dataset for the coterminous United States from landsat thematic mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67 (6), 650–661.

Wells, N., Goddard, S., Hayes, M.J., 2004. A self-calibrating palmer drought severity index. Journal of Climate 17 (12), 2335–2351.

Wilhite, D.A., 2002. Preparing for drought: a methodology. In: Wilhite, D.A. (Ed.), Drought: A Global Assessment, Hazards and Disaster Series, vol. II. Routledge, New York, pp. 89–104.

Yang, L., Wylie, B.K., Tieszen, L.L., Reed, B.C., 1998. An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the U.S. northern and central Great Plains. Remote Sensing of the Environment 65 (1), 25–37.