Determining agricultural drought using pattern recognition

V. Kumar, University of Manitoba
C.E. Haque, Brandon University
M. Pawlak, University of Manitoba

Abstract: Determining agricultural drought is a challenging task because of its complex and creeping nature. The deviation of crop yield from its long-term mean is applied to define agricultural drought. In the Canadian Prairies, nevertheless, no specific figure of such deviation has yet been established for this purpose. In this paper, we consider a case study of Swift Current, Saskatchewan. On the basis of varying percent deviation from the mean wheat yield, and using yield and climatic data (namely monthly temperature and precipitation for the period from 1975-1994), various possible scenarios of drought are examined. Data are divided into two distinct categories : i) drought and ii) no-drought. The error correction procedure of the pattern recognition is then applied in order to find a case wherein the drought could be linearly discriminated from the no-drought category. It is expected that this study will assist in defining and predicting drought in the Prairies.

Introduction

Agricultural drought refers to the significant reduction in crop yields due to soil moisture deficits. Due to an intricate relationship between crop yield and soil moisture deficit (Stewart 1983), analysis of agricultural drought becomes complex. The complexity is further increased by the difficulty in defining drought precisely (Yevjevich 1978). However, one of the common ways of defining drought is based on the deviation from the mean yield of a major crop in the region (Kumar 1993). A threshold value (percentage reduction from the mean yield) is selected and drought is considered to have occurred if crop yield is below the threshold.

In this paper, we choose different thresholds to define agricultural drought in Swift Current crop district, Saskatchewan, using yield of spring wheat, and climatic data. Subsequently, error-correction procedure of pattern recognition is applied to examine whether drought can be linearly classified. Success in classifying drought is expected to assist in drought prediction.

Pattern Recognition:

Pattern recognition (or machine perception) is a computerized process of classifying objects. Application of pattern recognition can be found in image processing, medical engineering, criminology (i.e., identification of finger prints), speech recognition, and signature identification (Duda and Hart 1973). In geography, pattern recognition can be applied in classifying patterns of some geographic phenomena, for example, drought. Process of pattern recognition begins with formation of pattern vectors which, in the present case, refer to drought vectors.

Vectorizing Drought:

Once drought is defined in terms of yield reduction, corresponding yield-affecting variables are chosen as elements of drought (or yield) vector. For example, if these elements are precipitation and average temperature, yield vectors can be shown as points in two-dimensional display (Figure 1).

As shown in the Figure 1, some of the yield vectors can be termed as drought vectors depending on the threshold value of the yield chosen to define drought. In the figure, there exist two categories: drought, and no-drought. Once the numbers of categories have been selected and corresponding vectors defined, error-correction procedure of pattern recognition can then be applied to test if a linear dividing line can be determined to separate the categories. In case of a positive result, the dividing line can be used to classify an unknown vector as belonging to one of the two categories.

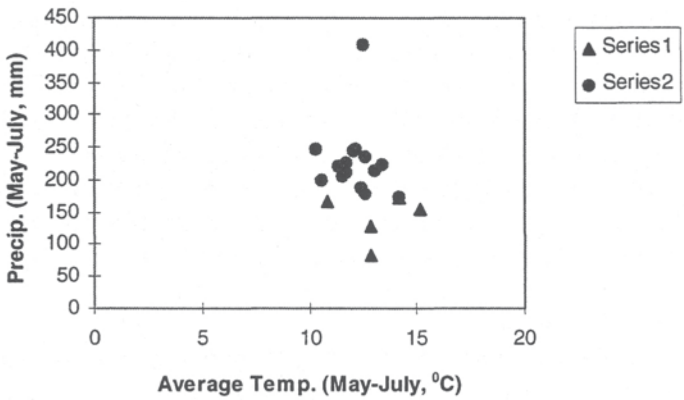

Figure 1: *Display of yield vectors, an example; Series 1: no drought, Series 2: drought.*

Error Correction Procedure:

The error-correction procedure is applied in three steps (Duda and Hart 1973) : i) All the vectors belonging to both the categories are augmented by adding additional element of 1; elements are the variables in different dimensions representing a vector, ii) All the elements in every vector of second class (say, drought vectors) are multiplied by -1 (Table 1) , and iii) Finally, a solution vector W is determined such that product of W with any yield vector exceeds zero i.e.,

$$\mathbf{Y_i W > 0} \qquad \textbf{for all i} \qquad [1]$$

where **i** is used to identify a vector and ranges from 1 to the total number of vectors in both categories.

The process of determining **W** commences with **W** as a unit vector i.e., all the elements of **W** are 1. It is followed by generating the product of **W** with individual yield vector one after the other. The moment the product does not satisfy the condition in case of any yield vector, the **W** is corrected as following:

$$\mathbf{W_{k+1} = W_k + c/k * Y_i,} \qquad [2]$$

Table 1: *Vectorizing yield elements; elements with negative sign represents drought.*

Year	Spring wheat yield	Percentage deviation from the	Drought identified by 'd' *	Elements in a yield vector		
	(t/ha)	mean yield		cv-t **	cv-p ***	Augmenting element
1975	1.67	-2.77		0.286	0.059	1
1976	2.21	28.67		0.170	0.198	1
1977	2.04	18.77		0.139	0.177	1
1978	1.71	-0.44		0.174	0.092	1
1979	1.53	-10.92	d	-0.302	-0.096	-1
1980	1.54	-10.34	d	-0.117	-0.161	-1
1981	1.8	4.80		0.260	0.160	1
1982	2.21	28.67		0.268	0.137	1
1983	1.88	9.46		0.308	0.200	1
1984	1.28	-25.48	d	-0.276	-0.154	-1
1985	0.91	-47.02	d	-0.188	-0.080	-1
1986	2.08	21.10		0.166	0.127	1
1987	1.86	8.29		0.142	0.124	1
1988	0.5	-70.89	d	-0.139	-0.098	-1
1989	1.67	-2.77		0.229	0.091	1
1990	2.08	21.10		0.223	0.152	1
1991	2.23	29.83		0.243	0.132	1

* when 10% deviation considered drought
** Coefficient of variation in temperature
*** Coefficient of variation in precipitation

where **c** is greater than 0 (chosen as 1 in the present case) and **k** starts with zero for unit vector **W** and is increased by one every time a correction is required. **Yi** is the yield vector causing correction in **W**. This process of correcting **W** continues until the required condition (Eq. 2) is met.

Methodology

In drought classification, selection of an appropriate definition of drought and subsequently the feature extraction to select elements to vectorize drought are essential components. Agricultural drought is qualitatively related to crop yield**,** but for a numerical analysis of drought, a quantitative definition is required. One way of defining drought quantitatively is through the use of an index crop i.e., a major crop in the region (Kumar 1993) . In the Canadian Prairies, yield of spring wheat, a major export crop, is therefore considered as a base to define drought.

After defining drought, variables are explored to characterize drought in form of drought vectors. Since drought is defined on the basis of yield, all the yield-affecting variables can be examined if they are to be considered as elements of drought vectors.

The crop yield depends on various factors (Parry et al. 1988). In the Prairies, however, weather is the main factor limiting yield and causing droughts (Walker 1989). The weather related variables that are consistently available across the Prairies are temperature and precipitation and they have been directly or indirectly used in yield estimation (Raddatz et al. 1994). Based on these considerations, temperature and precipitation data have been considered as basic parameters to form yield vectors. Additional variables are also derived from the basic parameters.

Data Collection and Analysis:

Considerations applied in defining and vectorizing drought led to the data collection. Spring wheat yield, monthly temperature, and precipitation were collected for the period 1975-1994 for the study area, Swift Current, Saskatchewan. Though we obtained the required data from the Canadian Wheat Board, their original sources were the Statistics of Canada, for yield data, and the Environment Canada, for weather data.

Various thresholds of percent deviation from the normal yield with an interval of 5% were considered with initial threshold of 10%. However, only thresholds of 10, 25, 45 and 70 were found to be relevant, other thresholds did not make a difference in the number of drought vectors defined on the basis of thresholds.

A total of 12 variables was considered to vectorize drought. Some of the basic variables are shown in the Table 2. The complete list of variables included i) Win-Tavg (average temperature from November to March), ii)Win-P (total precipitation from November to March), iii) May-T (average temperature in May), iv) May-P (total precipitation in May), Sum-Tavg (Average temperature from May to August), Sum-P (total precipitation from May to August), Sum1-Tavg (average temperature from May to July), Sum1-P (total

Table 2: *Some basic variables used to characterize yield or drought vectors.*

Year	yield (t/ha)	%dev mean yld	avg-t (nov-apr)	win-p (nov-apr)	may-t	may-p	jun-t	jun-p	jul-t	jul-p	aug-t	aug-p
1975	1.67	-2.77	-7.4	119.9	10.1	57.9	14.7	60.7	20.6	43.0	14.8	36.6
1976	2.21	28.67	-5.4	108.9	13.4	16.0	15.0	130.0	18.8	51.0	19.1	48.2
1977	2.04	18.77	-4.0	58.6	13.3	115.7	16.9	30.8	18.0	76.2	14.4	22.4
1978	1.71	-0.44	-8.9	140.5	12.15	53.85	17.1	49.65	18.5	34.45	17.5	21.7
1979	1.53	-10.92	-10.4	110.2	9.0	36.5	16.4	64.5	19.7	39.0	18.6	27.1
1980	1.54	-10.34	-5.1	68.1	14.1	12.4	16.2	94.4	18.4	66.3	15.0	42.9
1981	1.80	4.80	-3.0	85.7	12.1	26.3	12.95	103.8	18.6	79.8	20.6	29.3
1982	2.21	28.67	-8.4	109.4	9.0	121.2	15.6	32.1	17.7	102.8	17.0	51.2
1983	1.88	9.46	-4.6	117.4	9.6	45.8	16.1	27.1	19.1	121.7	21.3	28.9
1984	1.28	-25.48	-5.2	78.0	10.6	35.3	16.1	74.5	20.1	32.2	20.7	17.0
1985	0.91	-47.02	-7.1	103.4	13.8	34.1	13.7	16.1	20.1	20.7	16.9	26.6
1986	2.08	21.10	-5.6	84.4	12.2	97.3	17.2	54.7	17.3	68.9	18.1	22.8
1987	1.86	8.29	-2.1	69.3	14.1	22.6	18.6	31.9	18.3	72.6	14.8	57.1
1988	0.50	-70.89	-3.4	52.6	15.6	28.5	21.6	63.8	20.1	30.1	17.9	44.9
1989	1.67	-2.77	-7.1	107.4	11.5	77.3	16.4	90.2	20.5	32.4	18.1	74.6
1990	2.08	21.10	-3.9	109.3	10.8	52.2	16.5	62.8	17.8	102.3	18.7	15.9
1991	2.23	29.83	-5.2	132.7	11.4	71.7	16.0	152.3	18.7	89.5	20.8	41.4
1992	1.96		-2.3	73.0	11.2	41.8	15.4	34.6	16.4	72.2	16.1	49.0
1993	2.21		-5.4	108.7	12.5	15.2	14.3	80.6	15.6	86.9	16.3	112.7
1994	1.83		-6.9	116.7	12.2	65.8	15.7	76.2	18.9	29.3	18.1	37.4

precipitation from May to July), CV-T (coefficient of variation in temperature from May o August), CV-P (coefficient of variation in precipitation from May to August), CV1-T (coefficient of variation in temperature from May to July), CV1-P (coefficient of variation in precipitation from May to August).

Selecting two out of the total of 12, six pairs of variables were considered to vectorize yield. The pairs included i) Win-T, and Win-P, ii) May-T and May-P, iii) Sum-T and Sum-P, iv) Sum1-T and Sum1-P, v) CV-T and CV-P, and vi) CV1-T and CV1-P. From the four definitions of drought (thresholds 10, 25,45,70 %) and , for each drought definition , six definitions of yield vectors, a total of 24 cases were developed. The error-correction procedure was applied to investigate if the drought could be linearly classified. To accomplish this, a computer program was written to find out if, for any threshold or any definition of yield vector, it was possible to separate drought from the non-drought category.

Conclusion

In the classification procedure of the computer program, an iteration limit of 2000 was set. It was found that in no case did the solution vector W exist. Nevertheless, in the present study , only a limited number of cases were considered. Additional combinations need to be tested for a conclusive determination of drought characteristics. Further research in this direction is in progress. The use of non-linear techniques in classifying and predicting drought is being explored.

References

DUDA, R.O. and HART, P.E. 1973 *Pattern Classification and Scene Analysis* Toronto: John Wiley and Sons

KUMAR, V. 1993 'Predictive assessment of agricultural drought using climatic and remotely sensed data' Master of Engineering dissertation, University of Jodhpur, India

PARRY, M.L., CARTER, T.R. and KONIJN, N.T. ed. 1988 *Impact Of Climatic Variations On Agriculture Vol. 1, Assessment In Cool Temperature And Cold Regions* London: Kluwer Academic Publishers

RADDATZ, R.L., SHAYKEWICH, C.F. and BULLOCK, P.R. 1994 'Prairies crop yield estimates from modelled phenological development and water use' *Canadian Journal of Plant Science* 74: 429-436

SMART, G.M. 1983 'Drought analysis and soil moisture prediction' *J. Irrig. Drain. Div.* ASCE, 109(2):251-261

WALKER, G.K. 1989 'Model for operational forecasting of western Canada wheat yield' *Agricultural and Forest Meteorology* 44 :339-351

YEVJEVICH, V., HALL, W.A. and SALAS, J.D., ed. 1978 *Drought Research Needs* Colorado :Water Resource Publication, Fort Collins