Available online at www.sciencedirect.com

SCIENCE DIRECT®

Agricultural and Forest Meteorology 118 (2003) 49–62

AGRICULTURAL AND FOREST METEOROLOGY

www.elsevier.com/locate/agrformet

An evaluation of agricultural drought indices for the Canadian prairies

Steven M. Quiring [a,*], Timothy N. Papakryiakou [b]

[a] Department of Geography, Center for Climatic Research, University of Delaware, Newark, DE 19716, USA
[b] Department of Geography, Centre for Earth Observation Science, University of Manitoba, Winnipeg, MB, Canada R3T 2N2

Received 4 November 2002; received in revised form 20 March 2003; accepted 20 March 2003

Abstract

A comparative performance analysis was carried out on four agricultural drought indices to determine the most appropriate index for monitoring agricultural drought and predicting Canada Western Red Spring wheat (*Triticum aestivum* L.) yield on the Canadian prairies. A series of curvilinear regression-based crop yield models were generated for each of the 43 crop districts (20 in Saskatchewan, 12 in Manitoba, and 11 in Alberta) in the study region based on four commonly used measures of agricultural drought (Palmer Drought Severity Index, Palmer's *Z*-index, Standardized Precipitation Index, and NOAA Drought Index). The yield models were evaluated by comparing the model predicted yields to the observed yields (1961–1999) using four goodness-of-fit measures: the coefficient of determination (R^2), the index of agreement (d), the root mean square error (RMSE), and the mean absolute error (MAE). The analysis indicated that Palmer's *Z*-index is the most appropriate index for measuring agricultural drought in the Canadian prairies. The model evaluation indicated that the *Z*-index is best suited for predicting yield when there is significant moisture stress. There is a statistically significant relationship between the *Z*-index and Red Spring wheat yield in all crop districts, but the strength of the relationship varies significantly by crop district due to the influence of factors other than moisture availability (e.g. disease, pests, storm damage, and soil characteristics). The significant variations in model performance between the four agricultural drought indices underscores the necessity of carrying out a performance evaluation prior to selecting the most appropriate agricultural drought index for a particular application.

Keywords: Drought; Drought index; Wheat yield; *Triticum aestivum* L.

1. Introduction

Drought is a natural disaster that has a major economic impact on the Canadian prairies, owing to the vulnerability of the region's agricultural sector to weather variability. Recent growing season droughts in the Canadian prairies during 1988 and 2001 each resulted in more than 5 billion dollars in agricultural losses (Wheaton et al., 1992; Phillips, 2002).

While the effects of droughts are well documented, a proper, working definition of drought is less clear. Drought is a complex phenomenon that is difficult to accurately describe because its definition is both spatially variant and context dependent. Drought can be classified into three categories: meteorological, agricultural, and hydrological drought (Dracup et al., 1980). In this paper, we focus primarily on agricultural drought, which has been defined as an interval of time, generally of the order of months or years,

* Corresponding author. Tel.: +1-302-831-2344; fax: +1-302-831-6654.
E-mail address: squiring@udel.edu (S.M. Quiring).

0168-1923/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0168-1923(03)00072-8

when the moisture supply of a region consistently falls below the climatically appropriate moisture supply such that crop production or range productivity is adversely affected (adapted from Palmer, 1965; Rosenberg, 1978).

A drought index can be used to quantify: (1) the moisture condition of a region and thereby detect the onset and measure the severity of drought events, and (2) the spatial extent of a drought event thereby allowing a comparison of moisture supply conditions between regions (Alley, 1984). Drought indices can be useful tools for providing information for decision-makers in business, government and to the public stakeholders. These tools can be used to provide an early drought warning system (Lohani and Loganathan, 1997; Lohani et al., 1998), to calculate the probability of drought termination (Karl et al., 1987), to determine drought assistance (Wilhite et al., 1986), to assess forest fire hazard and dust storm frequency (Cohen et al., 1992; Wheaton, 1994), to predict crop yield (Sakamoto, 1978; Kumar and Panu, 1997), to examine the spatial and temporal characteristics of drought, the severity of drought, and to make comparisons between different regions (Alley, 1984, 1985; Soule, 1992; Kumar and Panu, 1997; Dai et al., 1998; Nkemdirim and Weber, 1999).

A large number of drought indices exist, each having a variety of data input requirements and each providing a somewhat different measure of drought. The more commonly used schemes include the Palmer Drought Severity Index (PDSI) and the Moisture Anomaly Index (Z-index) (Palmer, 1965), the Rainfall Anomaly Index (RAI) (van Rooy, 1965), the Crop Moisture Index (CMI) (Palmer, 1968), the Bhalme–Mooley Index (BMDI) (Bhalme and Mooley, 1980), the NOAA Drought Index (NDI) (Strommen et al., 1980; Titlow, 1987), the Standardized Anomaly Index (Katz and Glantz, 1986), the Standardized Precipitation Index (SPI) (McKee et al., 1993, 1995), Percent Normal, Deciles (Gibbs and Maher, 1967), and the Normalized Difference Vegetation Index-based Vegetation Condition Index (Kogan, 1995). Given the range in derivations and the different responses of these drought indices, not all are suitable for measuring agricultural drought.

The objective of this study is to compare the performance of the PDSI, Z-index, SPI, and NDI to determine which is the most appropriate for measuring agricultural drought in the Canadian prairies. These four drought indices have been chosen since each has been reported to quantify agricultural drought. The performance of the four drought indices are evaluated using four goodness-of-fit measures and possible sources of spatial variability in model performance are discussed.

2. Review of drought indices

2.1. Palmer drought severity index and moisture anomaly index

The PDSI and the Z-index were both developed by Palmer (1965) and have been widely used in the scientific literature (cf. Alley, 1984; Karl et al., 1987). The PDSI and Z-index are derived using a soil moisture/water balance algorithm that requires a time series of daily air temperature and precipitation data, and information on the available water content (AWC) of the soil. Soil moisture storage is handled by dividing the soil into two layers. The top layer has a field capacity of 25 mm, moisture is not transferred to the second layer until the top layer is saturated and runoff does not occur until both soil layers are saturated. Potential evapotranspiration (PE) is calculated using the Thornthwaite (1948) method and water is extracted from the soil by evapotranspiration when PE $>$ P (where P is the precipitation for the month). Evapotranspiration loss from the surface layer of the soil (L_s) always is assumed to take place at the potential rate. It is also assumed that the evapotranspiration loss from the underlying layer of the soil (L_u) depends on the initial moisture conditions in this layer, PE, and the combined available water content in both layers.

The Z-index is a measure of the monthly moisture anomaly and it reflects the departure of moisture conditions in a particular month from normal (or *climatically appropriate*) moisture conditions (Heim, 2002). The first step in calculating the monthly moisture status (Z-index) is to determine the expected evapotranspiration, runoff, soil moisture loss and recharge rates based on at least a 30-year time series. A water balance equation is subsequently applied to derive the expected, or normal precipitation. The monthly departure from normal moisture, d, is determined by comparing the expected precipitation to the actual precipitation. The Z-index, Z_i, then is the product of d and a weighting factor K for the month i,

$$Z_i = d_i K_i \tag{1}$$

where K_i is a weighting factor that is initially determined using an empirically derived coefficient, K', and then adjusted by a regional correction factor that is used to account for the variation between locations. Monthly values of K_i are calculated using

$$K_i = \frac{14.2}{\sum D_i K_i} K' \tag{2}$$

where D is obtained during the calibration period by determining the mean of the absolute values for each month of the year. In Eq. (2), a revised regional correction factor of 14.2, established by Akinremi et al. (1996), has been substituted for Palmer's original value of 17.67 (Palmer, 1965). Akinremi et al. (1996) found Palmer's original values artificially inflated the drought index values when applied to the Canadian prairies.

The PDSI, indicated by X_i, is a combination of Z_i, for the current month, and the PDSI value for the previous month,

$$X_i = \frac{Z_i}{3} + 0.897 X_{i-1} \tag{3}$$

While both the Z-index and the PDSI are derived using the same data, their monthly values are quite different. The Z-index is not affected by moisture conditions in the previous month, so Z-index values can vary dramatically from month to month. On the other hand, the PDSI varies more slowly because antecedent conditions account for two-thirds of its value. Although the PDSI was designed to measure meteorological drought (Palmer, 1965), it may be more appropriate as a measure of hydrological drought (Alley, 1985; Strommen and Motha, 1987; Akinremi et al., 1996) and, according to Karl (1986), the Z-index may be a better measure of meteorological or agricultural drought. It should be noted that although both the Z-index and PDSI are strongly weighted by both precipitation and temperature anomalies (Hu and Willson, 2000), the remaining indices are calculated using only precipitation. Alley (1984), Karl (1986), and Guttman (1998) have completed detailed evaluations of the limitations of the PDSI and Z-index, their work, along with the work of other researchers, has been summarized by Heim (2002).

2.2. NOAA drought index

The NDI was developed by Strommen and Motha (1987) as an early warning system for agricultural drought in developing countries. It was created to provide a simple and inexpensive means to monitor agricultural drought. The first step in calculating the NDI is to compute the mean monthly precipitation using at least 30 years of data (Titlow, 1987). Mean precipitation for each week of the year is derived and a running 8-week average of actual and average precipitation then is summed and compared. If actual precipitation is greater than 60% of the normal precipitation for the 8-week period, then the current week (week 8) is considered to have little or no water stress (Titlow, 1987). If any 8-week block receives less than 60% of normal precipitation, then water stress has occurred. Once stress is detected, weeks are no longer deleted when additional weeks are added until the moisture deficit is finally eliminated (e.g. actual precipitation exceeds 60% of normal) (Titlow, 1987). At that point, the evaluation period reverts to an 8-week block.

2.3. Standardized precipitation index

The SPI was developed by McKee et al. (1993, 1995) to provide a moisture supply index that performed better than the PDSI. The SPI is based on statistical probability and was designed to be a spatially invariant indicator of drought. It is produced by standardizing the probability of observed precipitation for any duration, durations of weeks or months can be used to apply this index to agricultural interests, and longer durations of years can be used to apply this index to water supply and water management interests (Guttman, 1999).

The SPI can be calculated for any location that has a long-term precipitation record. The precipitation record is fit with a probability density function (Pearson Type III) and subsequently transformed using an inverse normal (Gaussian) function (Guttman, 1998). This insures that the mean SPI value for any given location (and duration) is zero and the variance is one. Positive values of the SPI indicate greater than median precipitation, while negative values indicate

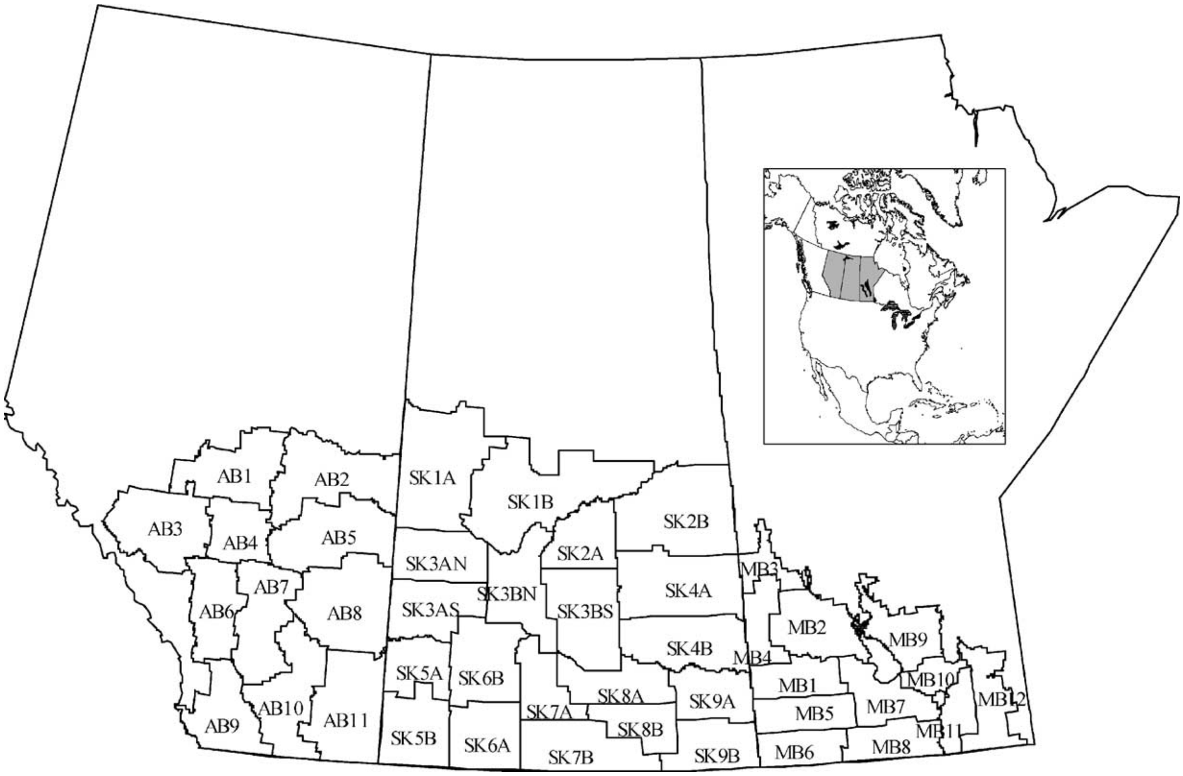


Fig. 1. The study regions—Alberta (AB), Saskatchewan (SK), and Manitoba (MB) crop districts. The study region is composed of a total of 43 crop districts (12 in Manitoba, 20 in Saskatchewan, and 11 in Alberta).

less than median precipitation. An SPI value of less than −1 indicates that a drought event is taking place and drought intensity can be calculated by summing the SPI values for all months within a drought event (McKee et al., 1993, 1995).

A standardized procedure was developed to compute the SPI so that all values are spatially and temporally comparable. If different probability distributions are used to calculate the SPI, different values for the index can be obtained. Guttman (1999) experimented with using different probability distributions and concluded that the Pearson Type III distribution provides the best model. However, determining the most effective distribution for evaluating monthly precipitation probabilities remains a matter for debate since Guttman's findings are contrary to the results of other researchers (cf. Legates, 1991). The calculation of all SPI values for this study will be carried out using Guttman's (1999) algorithm.

3. Methods

3.1. Study area

In this study, 43 crop districts from across the Canadian prairies (12 in Manitoba, 20 in Saskatchewan, and 11 in Alberta) were selected (Fig. 1). These crop districts are defined by the agricultural agencies in each province and are used to collect and disseminate agricultural data at the sub-provincial scale.

The Prairie ecozone covers an area of 520,000 km^2 across southern Manitoba, Saskatchewan, and Alberta (approximately 49–54°N latitude and 96–114°W longitude) (Raddatz, 1998). Annual precipitation in this zone ranges from 300 to 550 mm and the driest conditions in the prairies tend to be found in the south and the southwest, while the wettest areas are found in the north and northeastern prairies (Herrington et al., 1997). Approximately two-thirds of the precipitation

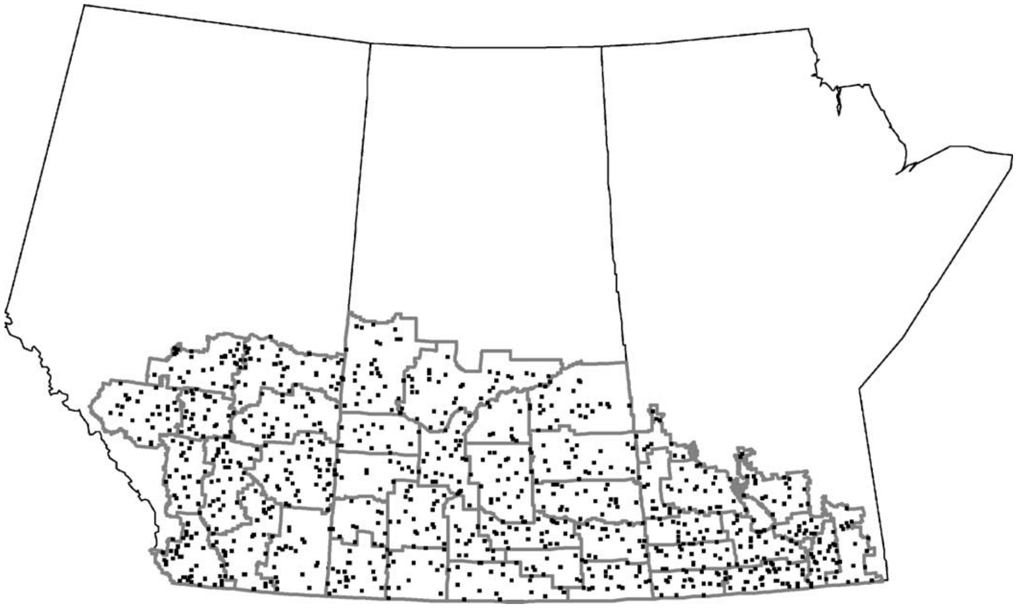


Fig. 2. Location of the climate stations providing temperature and/or precipitation data.

falls during the summer months (May to August) and a continuous snow cover lasts 4–5 months (Cohen et al., 1992).

3.2. Data

Daily air temperature and precipitation data were acquired from the Meteorological Service of Canada (MSC). Only stations with more than 5 years of data were used and the station list was also examined to insure that the stations included were evenly distributed within the crop districts. In total, data were acquired for a total of 1020 stations from across the prairies (Fig. 2). The stations were used to create a single daily time series of average temperature and precipitation (1920–1999) for each crop district. Some crop districts had a small number of days with missing data and these data were estimated by using the closest station from a neighbouring crop district. The PDSI, *Z*-index, SPI, and NDI were calculated for each of the 43 crop districts from 1920 to 1999 using the daily crop district time series of temperature and precipitation.

The AWC for each crop district was calculated using an area-weighted method from maps of Alberta, Saskatchewan and Manitoba that were developed by De Jong and Sheilds (1988). The latitude for each crop district was determined by calculating the centroid of each crop district polygon.

Canada Western Red Spring wheat (*Triticum aestivum* L.) yield data (hereafter referred to as Red Spring wheat) were available for 39 years (1961–1999) for Saskatchewan, 38 years for Alberta (1961–1998), and 23 years (1977–1999) for Manitoba crop districts. This crop was selected as a reference because it is grown extensively in all crop districts and because it accounts for almost 70% of the total wheat production in the region (Babb et al., 1997). The yield data were de-trended by regressing the average annual yield against the year-of-harvest for each crop district (Fig. 3). A positive trend in annual yield is due, in part, to farming innovations (Michaels, 1983; Babb et al., 1997). The resulting unstandardized residuals (hereafter referred to as yield departures) calculated for each crop district were used in the development and evaluation of yield models.

3.3. Evaluation of the agricultural drought indices

The evaluation of the agricultural drought indices was carried out in two stages. During the first stage, a growing season drought index variable was created

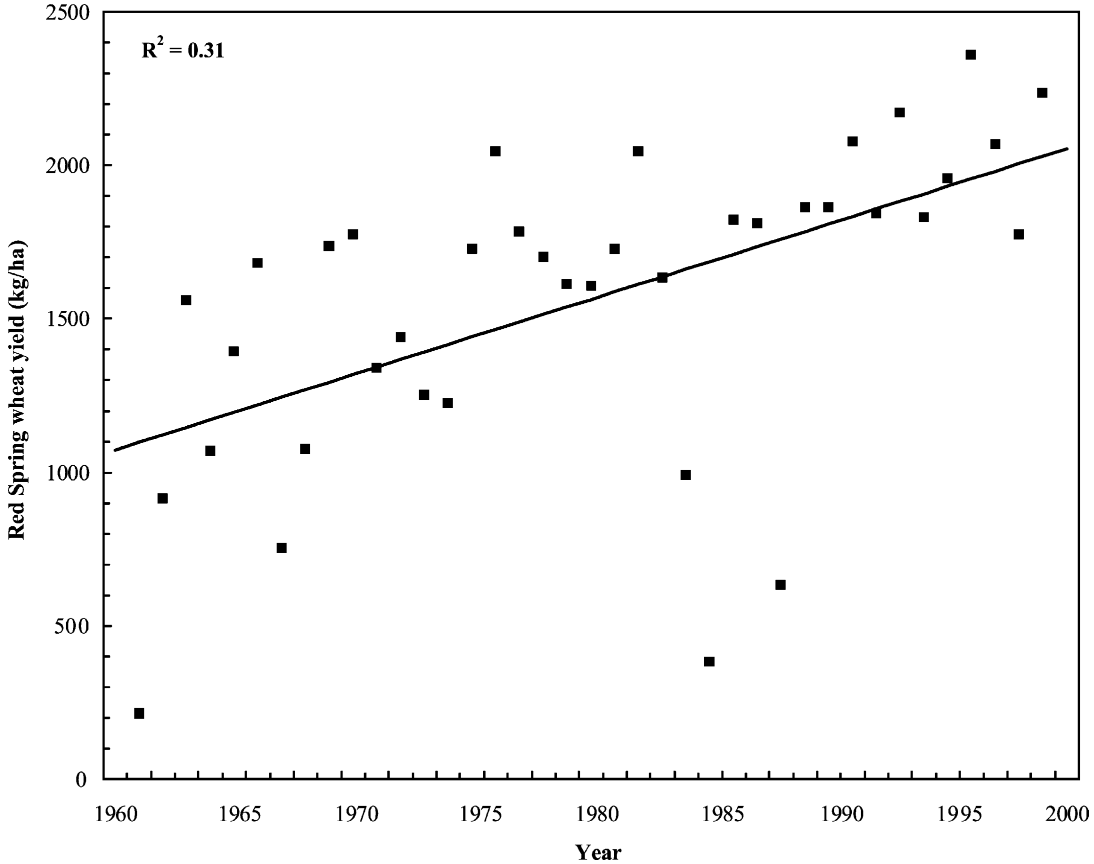


Fig. 3. Red Spring wheat yield (kg/ha) for crop district SK3BN (1961–1999). The solid line indicates the linear trend ($R^2 = 0.31$), which was removed from the yield data prior to analysis.

for each of the four drought indices by summing the monthly or weekly values of the indices for June and July. Although other combinations of drought index variables were tested, including those related to antecedent moisture conditions (during the winter/spring prior to planting) and late summer (August) moisture conditions, none of these combinations significantly improved the wheat yield models. These findings are similar to those of Arora et al. (1987), who determined that wheat yield is largely determined by moisture stress during the *heading* and *soft dough* stages (which usually occur during the later part of June and into July). The four growing season drought index variables were used to develop a set of curvilinear regression-based wheat yield models for each crop district/drought index combination (43 crop districts × 4 drought indices = 172 unique wheat yield models). Each model was derived using a second-order polynomial fit to the data, where the independent variable was a growing season drought index (either the PDSI, *Z*-index, SPI, or NDI) and the dependent variable was the yield departures for that crop district. Curvilinear regression (using a second-order polynomial) was used to determine the wheat yield models because it closely approximates the nature of the crop yield–water use relationship (Ash et al., 1992).

In the second stage, the wheat yield models were evaluated to determine the 'best' (most appropriate) agricultural drought index. The performance of the models was evaluated using four goodness-of-fit measures: the coefficient of determination (R^2), the root mean square error (RMSE), the mean absolute error (MAE), and the index of agreement (d). This approach provides an objective means of evaluating the drought indices that is consistent with our definition of agricultural drought and that has precedence in the literature (Akinremi et al., 1996; Kumar and Panu, 1997).

The coefficient of determination (R^2) describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0 to l, where the higher values indicate better agreement between the observed and predicted data. The interpretation of the coefficient of determination (R^2) is relatively straightforward (an R^2-value of 0.7 indicates that the model explains 70% of the variability in the observed data) and although, this statistic is frequently used, it is by no means a fail-proof method to evaluate model performance. Willmott (1984) and Legates and McCabe (1999) indicate that R- and R^2-values are inappropriate statistics for quantitative comparisons of model performance because they are highly sensitive to outliers. However, because the coefficient of determination is commonly used, we have chosen to include it along with two measures of error (RMSE and MAE) and one measure of model fit (d) to evaluate model performance.

The root mean square error is calculated using

$$\mathrm{RMSE} = \sqrt{N^{-1}\sum_{i=1}^{N}(O_i - P_i)^2} \tag{4}$$

and the mean absolute error is calculated using

$$\mathrm{MAE} = N^{-1}\sum_{i=1}^{n}|O_i - P_i| \tag{5}$$

The RMSE and MAE are both error measures used to represent the average differences between model predicted (P) and observed (O) values. It is important to include absolute error measures (such as MAE and RMSE) in a model evaluation because they provide an estimate of model error in the units of the variable (e.g. kg/ha) (Legates and McCabe, 1999). The MAE provides a more robust measure of average model error than the RMSE, since it is not influenced by extreme outliers (Legates and McCabe, 1999).

The index of agreement (d) is calculated using

$$d = 1.0 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(|Pi - \bar{O}| + |O_i - \bar{O}|)^2} = 1.0 - N\frac{\mathrm{MSE}}{\mathrm{PE}} \tag{6}$$

The index of agreement is the ratio between the mean square error (MSE) and the potential error (PE) and it measures the degree to which the observed data are approached by the predicted data. The index of agreement varies between 0 and 1, where 0 indicates no agreement between the predicted and observed data and 1 indicates perfect agreement ($O_i = P_i$ for all data pairs). Thus, this statistic measures the deviations from the 1:1 line (the line that results if the observed and predicted data are equivalent). Since d is dimensionless, it is easier to interpret than other measures of relative difference and it overcomes the insensitivity of correlation-based measures (e.g. R^2) to differences in the observed and predicted means and variances (Willmott, 1984). One of the problems with the index of agreement is that only values of 0 and 1 have a clear physical meaning.

Overall, no single model evaluation statistic is sufficient to evaluate model performance because, as described above, each has unique strengths and weaknesses. Therefore, applying a range of evaluation statistics provides a more solid basis for understanding model performance and undertaking model comparisons (Willmott, 1984).

4. Results and discussion

4.1. Model evaluation

The mean model performance statistics for all 43 crop districts are summarized in Table 1. Of the four agricultural drought indices, the Z-index ranked first in all four goodness-of-fit measures. The Z-index correlates best with the yield departures ($R^2 = 0.47$), it agrees most closely with the observed data ($d = 0.76$), and it has the least amount of error (RMSE and MAE values of 256.04 and 206.84 kg/ha, respectively). The SPI ranked second in all four model-performance

Table 1
Mean model performance statistics for all crop districts

	R^2	d	RMSE	MAE
Z model	0.47	0.76	256.04	206.84
SPI model	0.33	0.63	295.28	234.40
PDSI model	0.27	0.56	306.38	255.49
NDI model	0.15	0.41	333.97	269.86

statistics, followed by the PDSI. The NDI is the least effective index for monitoring agricultural drought and it appears to have only a weak relationship with yield departures.

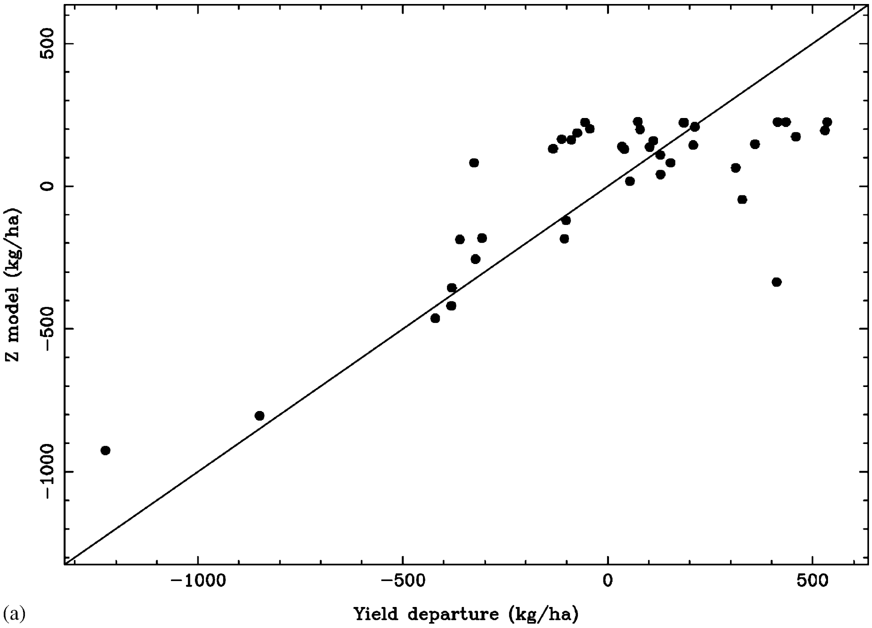
The performance evaluation graphic shown in Fig. 4 can be used to illustrate model performance and to diagnose model bias. Fig. 4 shows the predicted (modeled) versus observed yield departures (kg/ha) for the four drought-index-based models from a representative (one that is close to average for all of the goodness-of-fit measures) crop district (SK6A). If any of the models were able to perfectly predict yield departures, all of the points would fall along the 1:1 line (the perfect prediction line). Points that are above this line indicate that the model is over-predicting yield departures (predicting higher yields than are actually received), and points falling below the line indicate that the model is under-predicting yield (predicting lower yields than are actually received).

The yield model based on the *Z*-index (Fig. 4a) does reasonably well at predicting yields during those years when yield departures are strongly negative to moderately positive (seasons having yield departures between −1000 and +200 kg/ha). However, the model appears to systematically under-predict yield departures during the years with yield departures greater than +200 kg/ha. The yield models based on the SPI (Fig. 4b) and PDSI (Fig. 4c) both exhibit more scatter around the perfect prediction line than the *Z*-index. Both the SPI and PDSI also systematically over-predict during years observing negative yield departures and systematically under-predict during years observing positive yield departures. There appears to be very little agreement between the yield departures predicted by the NDI and the observed values (Fig. 4d).

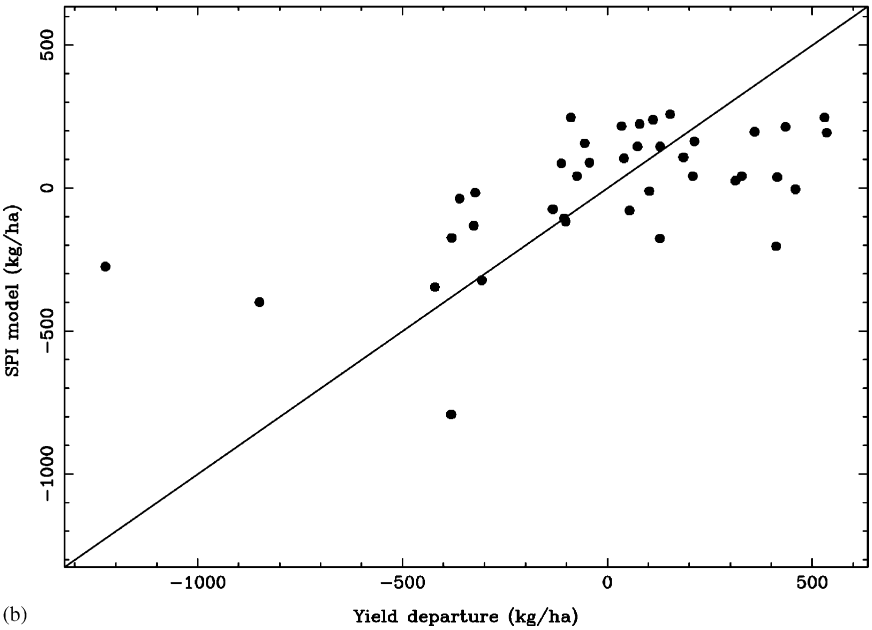
Fig. 4 shows that the *Z*-index has the least scatter around the perfect prediction line and therefore it is the most appropriate drought index for predicting yield departures. However, although the *Z*-index does reasonably well at predicting yield departures during most years (Fig. 5), the model has a systematic bias that causes it to under-estimate yield departures during years with large yields. In Fig. 5, the model predicted yield departures tend to agree most closely with the observed yield during those years that that have large negative values (e.g. 1961, 1967, and 1984). There is much less agreement between the predicted and observed yield departures during years that have large positive values (e.g. 1963, 1976, and 1982). This suggests that the *Z*-index-based yield model is best suited for predicting yield departures during those growing seasons where moisture stress occurs and that it does not work as well during years with optimum moisture conditions (e.g. moisture is not a limiting factor). While it may be possible to revise the yield model based on the *Z*-index to remove some of the systematic model bias, there is still a great deal of unsystematic (random) bias. When moisture conditions are not limiting, other factors (e.g. soil fertility, the presence or absence of disease/pests, and the amount of fertilizer applied) become important determinants of crop yield.

4.2. Discussion of model results

It is not surprising that the *Z*-index proved to be the best index for measuring agricultural drought on the Canadian prairies since the *Z*-index is more physically-based than the other drought indices. The *Z*-index requires daily temperature and precipitation data and the AWC of the soil. In addition, unlike the SPI and NDI, both of the Palmer drought indices (*Z*-index and PDSI) also consider potential and actual evapotranspiration, important variables in determining crop growth. Despite the similarities in derivation and the common use of the PDSI for measuring agricultural drought, the *Z*-index clearly outperformed the PDSI in this study. This occurred because the PDSI is highly dependent on antecedent conditions (e.g. it has a long 'memory') (Guttman, 1998). The majority of the value of the PDSI in any given month is determined by antecedent conditions (previous PDSI values), only one-third of its value is dependent on the temperature and precipitation in the current month. Since crop growth is highly dependent on short-term moisture conditions, the *Z*-index is more appropriate than the PDSI for agriculture and forestry applications because it is more responsive to those conditions (Karl, 1986). Akinremi et al. (1996) attempted to correct some of the limitations of the PDSI by coupling it with the Versatile Soil Moisture Budget to better represent the soil moisture balance. While this approach significantly improved the ability of the PDSI to measure agricultural drought it cannot be applied in this study since it was designed for use

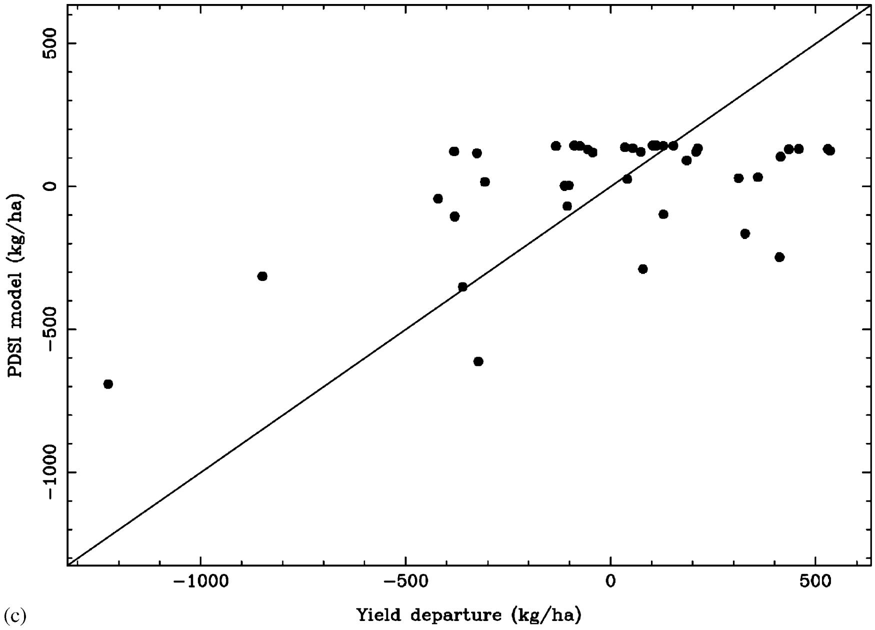


(a)

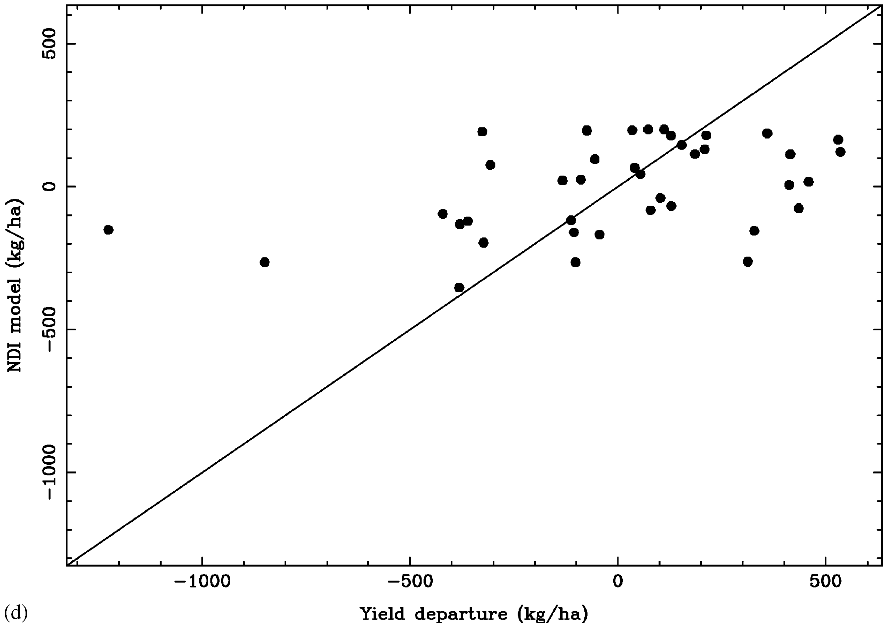


(b)

Fig. 4. Predicted (modeled) versus observed Red Spring wheat yield departures (kg/ha) for crop district SK6A: (a) *Z*-index, (b) SPI, (c) PDSI, and (d) NDI. The solid line denotes the perfect prediction line (if the model were able to perfectly predict yield all of the points would fall along this line). Points that are above this line indicate that the model is over-predicting yield (predicting higher yields than are actually received), and points below the line indicate that the model is under-predicting yield (predicting lower yields than are actually received).



(c)



(d)

Fig. 4. (*Continued*).

at a specific site, not over a large region (e.g. crop district). The Versatile Soil Moisture Budget also requires detailed site-specific data (e.g. initial moisture conditions, date of occurrence for five phonological stages, the depth and water-holding characteristics of each layer of the soil, etc.) (Akinremi et al., 1996).

The SPI, although it did not perform as well as the *Z*-index, is also a useful measure of agricultural

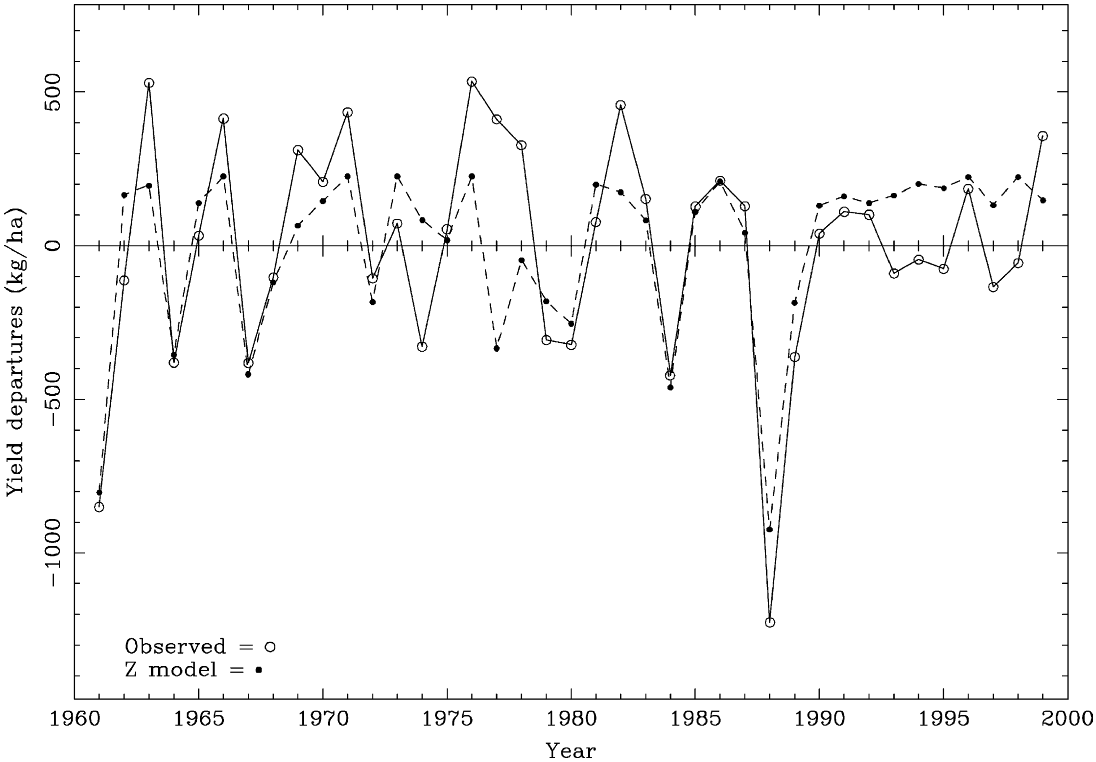


Fig. 5. Time series (1961–1999) of yield departures predicted using *Z*-index model (filled circle) and observed Red Spring wheat yield departures (open circle) for crop district SK6A. The solid line connects the observed yield values and the dashed line connects the model predicted values.

drought. The SPI has a number of important advantages over the *Z*-index: it is simpler to calculate, easier to understand, spatially invariant, and it can be calculated for any period of interest (Guttman, 1998). Despite the fact that the SPI relies solely on monthly precipitation values, it outperformed the PDSI in all four of the model performance statistics. The NDI is also calculated based solely on precipitation, but the arbitrary thresholds used by the NDI for determining the onset and termination of drought events (e.g. drought occurs when precipitation is less than 60% of normal) obviously render it unsuitable for predicting crop yields.

4.3. Spatial variability in model performance

Table 2 illustrates the considerable spatial variability in model performance that is present across the study region. The *Z*-index model performed the best in SK3AS ($R = 0.88$) and performed the worst in AB9 ($R = 0.29$). Therefore, while the *Z*-index model was able to explain approximately 78% of variance in the yield departure data in SK3AS, it was able to explain only approximately 9% of variance in AB9. The variability of model of performance can also be illustrated by looking at the distribution of index of agreement (d)-values for the *Z*-index (Fig. 6). The majority of crop districts (33 out of 43) have a d-value greater than 0.70, indicating that the *Z*-index does a reasonable job of predicting yield departures in most crop districts.

These differences in model performance from one crop district to another may occur due to regional variations in climate, soil conditions, response to moisture stress, or data quality. In addition, although Red Spring wheat yield is sensitive to growing season moisture conditions, there many other factors, such as disease, pests, storm damage (both wind and hail), and soil characteristics, that affect yield. The distribution of

Table 2
Z-index model statistics for all 43 crop districts

Crop district	*N*	*R*	*F*
AB1	38	0.79	58.3
AB2	38	0.67	29.9
AB3	38	0.56	16.3
AB4	38	0.64	24.3
AB5	38	0.70	33.8
AB6	38	0.72	38.2
AB7	38	0.70	34.0
AB8	38	0.30*	3.6*
AB9	38	0.29*	3.3*
AB10	38	0.70	34.8
AB11	38	0.31*	3.8*
SK1A	39	0.66	28.1
SK1B	39	0.75	48.4
SK2A	37	0.86	97.3
SK2B	39	0.70	34.8
SK3AN	39	0.84	87.0
SK3AS	39	0.88	127.0
SK3BN	39	0.77	54.0
SK3BS	39	0.78	55.9
SK4A	39	0.83	82.3
SK4B	39	0.75	48.6
SK5A	39	0.76	49.4
SK5B	39	0.75	46.2
SK6A	39	0.78	56.9
SK6B	39	0.82	76.3
SK7A	39	0.70	35.9
SK7B	39	0.73	43.0
SK8A	39	0.57	17.8
SK8B	39	0.66	28.8
SK9A	39	0.68	32.5
SK9B	39	0.63	24.2
MB1	23	0.68	17.6
MB2	23	0.76	29.5
MB3	23	0.73	24.3
MB4	23	0.45*	5.5*
MB5	23	0.58	10.4
MB6	23	0.59	11.3
MB7	23	0.80	38.2
MB8	23	0.81	41.3
MB9	23	0.53	8.1
MB10	23	0.65	15.0
MB11	23	0.70	20.1
MB12	23	0.48	6.4

The Pearson correlation coefficients (*R*)- and *F*-values have been tested ($\alpha = 0.01$) for significance. All of the *R*- and *F*-values are statistically significant at $\alpha = 0.01$ expect for those with a (*), they are statistically significant at $\alpha = 0.05$.

precipitation during the summer and its relationship to the phenological development of Red Spring wheat is another important factor that is not currently accounted for in the models. As discussed earlier, it appears that

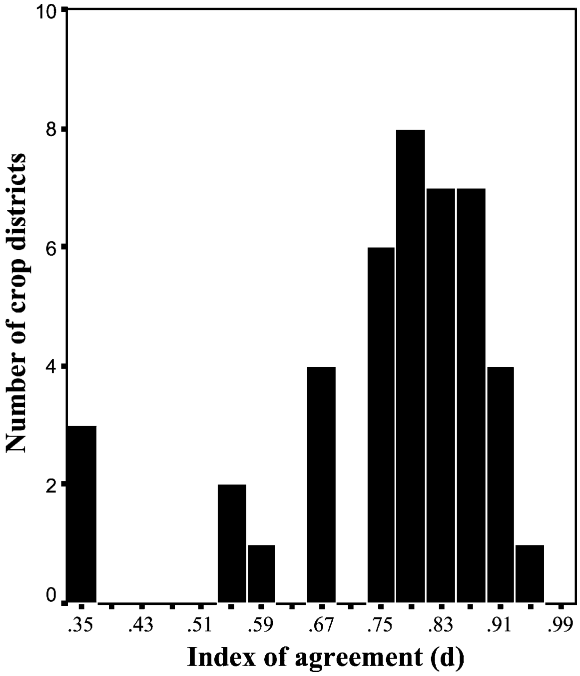


Fig. 6. The histogram shows the distribution of index of agreement (*d*)- values obtained by using the *Z*-index model. The majority of crop districts (33 out of 43) have a *d*-value >0.70.

the yield models based on the *Z*-index are best suited for predicting yield departures during growing seasons where significant moisture stress occurs. Thus, some of the spatial variability in model performance may be related to variations in mean growing season precipitation. The *Z*-index-based yield models will tend to perform better in those areas that typically receive less precipitation during the growing season.

Overall, while there is a great deal of variability across the study region, there was a statistically significant correlation between the *Z*-index and yield departures in all of the crop districts (39 crop districts at $\alpha = 0.01$, and 4 crop districts at $\alpha = 0.05$).

5. Conclusions

The analysis of the four agricultural drought indices indicated that the *Z*-index is the most appropriate for measuring agricultural drought in the Canadian prairies. The model evaluation indicated that the *Z*-index is best suited for predicting yield during growing seasons (and in locations) where significant moisture stress occurs. There is a statistically significant relationship between the *Z*-index and Red Spring wheat yield in all crop districts, but the strength of the relationship varies significantly by crop district.

The major differences in the ability of the four measures of agricultural drought to predict crop yield illustrate why it is so important to select an appropriate agricultural drought index. Regardless of whether the agricultural drought index will be used for quantifying/monitoring agricultural drought, predicting yield, or determining drought insurance/drought assistance, the decision on which drought index to use will have a major impact on the accuracy of your information. Choosing the most appropriate measure of agricultural drought is particularly difficult because the answer will vary depending on the crop, the study region, and the spatial scale of the intended application. Therefore, a variety of indices should always be tested to determine the most appropriate drought index for a particular application.

Acknowledgements

S.M. Quiring was partially supported by a Post-Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank N. Guttman for providing assistance with the SPI code and D.R. Legates for his constructive comments on the manuscript. Helpful comments from the anonymous reviewers were greatly appreciated.

References

Akinremi, O.O., McGinn, S.M., Barr, A.G., 1996. Evaluation of the Palmer Drought Index on the Canadian prairies. J. Climate 9, 897–905.

Alley, W.M., 1984. The Palmer Drought Severity Index: limitation and assumptions. J. Climate Appl. Meteor. 23, 1100–1109.

Alley, W.M., 1985. The Palmer Drought Severity Index as a measure of hydrologic drought. Water Resour. Bull. 21 (1), 105–114.

Arora, V.K., Prihar, S.S., Gajri, P.R., 1987. Synthesis of a simplified water use simulation model for predicting wheat yields. Water Resour. Res. 23, 903–910.

Ash, G.H.B., Shaykewich, C.F., Raddatz, R.L., 1992. Moisture risk assessment for spring wheat on the eastern prairies: a water-use simulation model. Climatol. Bull. 26 (2), 65–78.

Babb, J.C., Khandekar, M.L., Garnett, E.R., 1997. An analysis of PNA indices for forecasting summer weather over the Canadian prairies with implications for wheat yield and protein. In: Proceedings of the Long-Range Weather and Crop Forecasting Work Group Meeting III. Canadian Meteorological Centre, Dorval, QB, pp. 31–36.

Bhalme, H.N., Mooley, D.A., 1980. Large-scale droughts/floods and monsoon circulation. Mon. Wea. Rev. 108, 1197–1211.

Cohen, S., Wheaton, E.E., Masterton, J., 1992. Impacts of Climatic Change Scenarios in the Prairie Provinces: A Case Study from Canada. Canadian Climate Center, Downsview, ON.

Dai, A., Trenberth, K.E., Karl, T.R., 1998. Global variations in droughts and wet spells: 1900–1995. Geophys. Res. Lett. 25 (17), 3367–3370.

De Jong, R., Sheilds, J.A., 1988. Available water-holding capacity maps of Alberta, Saskatchewan, and Manitoba. Can. J. Soil Sci. 68, 157–163.

Dracup, J.A., Lee, K.S., Paulson, E.G.J., 1980. On the definition of drought. Water Resour. Res. 16 (2), 297–302.

Gibbs, W.J., Maher, J.V., 1967. Rainfall Deciles as Drought Indicators. Bureau of Meteorology Bulletin, No. 48, Commonwealth of Australia, Melbourne.

Guttman, N.B., 1998. Comparing the Palmer Drought Index and the Standardized Precipitation Index. J. Am. Water Resour. Assoc. 34 (1), 113–121.

Guttman, N.B., 1999. Accepting the Standardized Precipitation Index: a calculation algorithm. J. Am. Water Resour. Assoc. 35 (2), 311–322.

Heim Jr, R.R., 2002. A review of twentieth-century drought indices used in the United States. Bull. Am. Meteor. Soc. 83 (8), 1149–1165.

Herrington, R., Johnson, R., Hunter, F., 1997. Responding to global climate change in the Canadian prairies. In: Canada Country Study: Climate Impacts and Adaptation, vol. III. Environment Canada, Ottawa, ON.

Hu, Q., Willson, G.D., 2000. Effects of temperature anomalies on the Palmer Drought Severity Index in the central United State. Int. J. Climatol. 20, 1899–1911.

Karl, T., 1986. The sensitivity of the Palmer Drought Severity Index and Palmer's *Z*-Index to their calibration coefficients including potential evapotranspiration. J. Climate Appl. Meteor. 25, 77–86.

Karl, T., Quinlan, F., Ezell, D.S., 1987. Drought termination and amelioration: its climatological probability. J. Climate Appl. Meteor. 26, 1198–1209.

Katz, R.W., Glantz, M.H., 1986. Anatomy of a rainfall index. Mon. Wea. Rev. 114, 764–771.

Kogan, F.N., 1995. Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. Bull. Am. Meteor. Soc. 76 (5), 655–668.

Kumar, V., Panu, U., 1997. Predictive assessment of severity of agricultural droughts based on agro-climatic factors. J. Am. Water Resour. Assoc. 33 (6), 1255–1264.

Legates, D.R., 1991. An evaluation of procedures to estimate monthly precipitation probabilities. J. Hydrol. 122, 129–140.

Legates, D.R., McCabe, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35 (1), 233–241.

Lohani, V.K., Loganathan, G.V., 1997. An early warning system for drought management using the Palmer Drought Index. J. Am. Water Resour. Assoc. 33 (6), 1375–1386.

Lohani, V.K., Loganathan, G.V., Mostaghimi, S., 1998. Long-term analysis and short-term forecasting of dry spells by Palmer Drought Severity Index. Nordic Hydrol. 29 (1), 21–40.

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th Conference on Applied Climatology. AMS, Boston, MA, pp. 179–184.

McKee, T.B., Doesken, N.J., Kleist, J., 1995. Drought monitoring with multiple time scales. In: Proceedings of the 9th Conference on Applied Climatology. AMS, Boston, MA, pp. 233–236.

Michaels, P.J., 1983. Price, weather, and "acerage abandonment" in Western Great Plains wheat culture. J. Climate Appl. Meteor. 22, 1296–1303.

Nkemdirim, L., Weber, L., 1999. Comparison between the droughts of the 1930s and the 1980s in the Southern Prairies of Canada. J. Climate 12, 2434–2450.

Palmer, W.C., 1965. Meteorological Drought. Research Paper No. 45, US Weather Bureau, Washington, DC.

Palmer, W.C., 1968. Keeping track of crop moisture conditions, nationwide: the Crop Moisture Index. Weatherwise 21, 156–161.

Phillips, D., 2002. The top ten Canadian weather stories for 2001. CMOS Bull. 30 (1), 19–23.

Raddatz, R.L., 1998. Anthropogenic vegetation transformation and the potential for deep convection on the Canadian prairies. Can. J. Soil Sci. 78, 657–666.

Rosenberg, N.J., 1978. North American Droughts. AAAS, Boulder, CO.

Sakamoto, C.M., 1978. The *Z*-index as a variable for crop yield estimation. Agric. Meteor. 19, 305–313.

Soule, P.T., 1992. Spatial patterns of drought frequency and duration in the contiguous USA based on multiple drought event definitions. Int. J. Climatol. 12, 11–24.

Strommen, N., Krumpe, P., Reid, M., Steyaert, L., 1980. Early warning assessments of droughts used by the U.S. agency for international development. In: Pocinki, L.S., Greeley, R.S., Slater, L. (Eds.), Climate and Risk. The MITRE Corporation, McLean, VA, pp. 8–37.

Strommen, N.D., Motha, R.P., 1987. An operational early warning agricultural weather system. In: Wilhite, D.A., Easterling, W.E., Wood, D.A. (Eds.), Planning For Drought: Toward a Reduction of Societal Vulnerability. Westview Press, Boulder, CO.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev. 38, 55–94.

Titlow, J.K., 1987. A precipitation-based drought index for the Delaware river basin. Publications in Climatology 40. C.W. Thornthwaite Associates, Centerton, NJ.

van Rooy, M.P., 1965. A rainfall anomaly index independent of time and space. Notos 14, 43–48.

Wheaton, E.E., 1994. Impacts of a variable and changing climate on the Canadian prairies provinces: A preliminary intergration and annotated bibliography. SRC Publication No. E-2900-7-E-93, Saskatchewan Research Council, Saskatoon, SK.

Wheaton, E.E., Arthur, L.M., Chorney, B., Shewchuk, C., Thorpe, J., Whitting, J., Whittrock, K., 1992. The prairie drought of 1988. Climatol. Bull. 26, 188–205.

Wilhite, D.A., Rosenberg, N.J., Glantz, M.H., 1986. Improving federal reponse to drought. J. Climate Appl. Meteor. 25, 332–342.

Willmott, C.J., 1984. On the evaluation of model performance in physical geography. In: G.L. Gaile, C.J. Willmott (Eds.), Spatial Statistics and Models. D. Reidel, Norwell, MA.