INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2011)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2385

Multiple drought indices for agricultural drought risk assessment on the Canadian prairies

Liu Sun,[a,b]* Scott W. Mitchell[a,b] and Andrew Davidson[a,c]

[a] *Department of Geography and Environmental Studies, Carleton University, Ottawa, Ontario K1S 5B6, Canada*
[b] *Geomatics and Landscape Ecology Research Laboratory, Carleton University, Ottawa, Ontario K1S 5B6, Canada*
[c] *National Land and Water Information Service, Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0K9, Canada*

ABSTRACT: A multi-index drought (MID) model was developed to combine the strengths of various drought indices for agricultural drought risk assessment on the Canadian prairies, as related to spring wheat crop yield. The model automatically selects and combines optimum drought indices derived from the preceding and current months as they become available to better match the conditions (both spatially and temporally) where they work well. The cross-validation results showed that (1) the prediction accuracy of the MID model is better than (or occasionally equal to) using any single drought index for all modelling stages, (2) drought indices derived from the recharge period are useful for early drought risk detection, (3) model prediction accuracy improved as the growing season progressed with the most accurate assessments at the beginning of August, and (4) the model performed best in the more arid locations in the southern prairies, which tend to have a more variable precipitation regime. The model assessment results provide the spatial intensity distribution of possible drought progression and recession before and during the growing season, and can be used with complementary information in agricultural drought risk management and mitigation strategies. Copyright © 2011 Royal Meteorological Society

KEY WORDS agricultural drought; Canadian prairies; drought index; spring wheat yield; risk assessment

Received 8 July 2010; Revised 10 March 2011; Accepted 28 May 2011

1. Introduction

Drought is a recurrent phenomenon on the Canadian prairies (hereafter referred to as 'the prairies'). Three decades in the twentieth century (1910–1919, 1930–1939, and 1980–1989) experienced drought for more than half the decade (Nkemdirim and Weber, 1999). Droughts on the prairies are closely related to the lack of precipitation, above-normal air temperatures, low soil moisture, and insufficient surface water supply (Wheaton *et al*., 1992).

Although the natural conditions of the prairies are favourable for mechanized farming, the agricultural sector is highly vulnerable to weather variability (Quiring and Papakyriakou, 2003). Recent growing season droughts in the prairies during 2001 and 2002 resulted in an estimated loss of $3.6 billion in agricultural production (Wheaton *et al*., 2005). To reduce the serious consequences of drought, besides improving the understanding of the hazard and the factors that influence vulnerability, there are calls for more attention to prediction/early warning activities (i.e. risk assessment) that could improve drought preparedness and response, as well as to reduce future impacts (Sivakumar and Wilhite, 2002).

* Correspondence to: L. Sun, c/o Scott Mitchell, Department of Geography and Environmental Studies, Carleton University, B349 Loeb Building, 1125 Colonel By Drive, Ottawa, ON, K1S 5B6, Canada. E-mail: lsun5@connect.carleton.ca

Drought indices are the integration of one or more climate or hydrological variables (e.g. precipitation, temperature, soil moisture, stream flow, groundwater levels, etc.) on a quantitative scale (Steinemann *et al*., 2005). They have been widely used to detect the onset and severity of drought, and to study its spatial and temporal patterns (Quiring and Papakyriakou, 2003). The relatively low cost and generally high availability of the weather and hydrological data compared to other types of data are other reasons for their wide application.

Over the years, to our knowledge, the majority of drought indices studies have focused on evaluating drought indices for specific regions (e.g. Quiring and Papakyriakou, 2003; Morid *et al*., 2006; Mavromatis, 2007) or the use of a single well-developed drought index to characterize and predict droughts over specific regions (e.g. Tsakiris and Vangelis, 2004; Cancelliere *et al*., 2007). However, because each index provides a somewhat different measure of drought, use of a particular specific index has often been demonstrated to be inadequate for completely representing this complex phenomenon (Heim, 2002; Steinemann *et al*., 2005; Quiring, 2009). A combination of various drought indices may provide a more comprehensive assessment of drought conditions than a single-index approach, but this has been challenging because there has been a lack of systematic methods for their combination, use, and evaluation (Steinemann and Cavalcanti, 2006).

In recent years, significant progress has been made in employing multiple drought indices in drought management. For example, the Objective Blend of Drought Indicators (OBDI) was developed for the U.S. Drought Monitor to provide a comprehensive assessment of drought conditions across the country (Svoboda *et al.*, 2002). The OBDI combines three climate-based drought indices and the Climate Prediction Center's soil moisture model. A percentile approach was used to transform all input data into a standardized scale to which drought category thresholds and weights for each individual index could be assigned. The OBDI is beneficial in providing a single 'average' drought designation at the national and state level, but is not meant to capture local drought conditions, such as in individual counties (Steinemann *et al.*, 2005). Brown *et al.* (2008) developed a hybrid geospatial drought monitoring tool, the Vegetation Drought Response Index (VegDRI), to produce a near real-time 1 km resolution map of drought conditions in seven north-central states of the United States. The model integrates climate-based drought indices, satellite-derived vegetation condition information, and other biophysical information. The model was empirically derived for three seasonal phases (spring, summer, and fall) by applying a supervised classification and regression tree analysis for each phase. The VegDRI map provides more localized drought information at a county to sub-county scale.

The purpose of this article is to present an operational model framework that combines the strengths of various drought indices to provide a more comprehensive assessment of agricultural drought conditions in the Canadian prairies. Agricultural drought has been defined as 'the condition when moisture supply of a region consistently fails to meet the needs of a particular crop at a particular time, such that the crop production or range productivity is significantly affected' (Bordi and Sutera, 2007). As agricultural drought on the prairies is the single most limiting factor to crop yield (Akinremi *et al.*, 1996), it can be used as an agricultural drought indicator: By predicting reduced crop yield, one can predict droughts (Morgan, 1985; Sinha *et al.*, 1992). As drought index integration is the main focus of this study, other management (e.g. soil fertility status, cultivation practices, pest control, and crop disease) and weather-related factors that effect crop yield were not considered and were assumed to remain stable throughout the study period. To better match crop phenological stages and to detect short-term dry spells, especially at pre-planting and early crop growth stages, agricultural drought risk is assessed at pre-planting and at the beginning of each month during the growing season. The risk assessment results are mapped monthly to provide the spatial intensity distribution of possible drought progression and recession in the prairies.

2. Materials and methods

2.1. Study area

The study area is composed of a total of 34 Census Agricultural Regions (CARs) across the prairies (Figure 1). These CARs are composed of groups of adjacent census divisions, which were defined by the agricultural agencies in each province for the purposes of regional planning and managing common services (Statistics Canada, 2003).

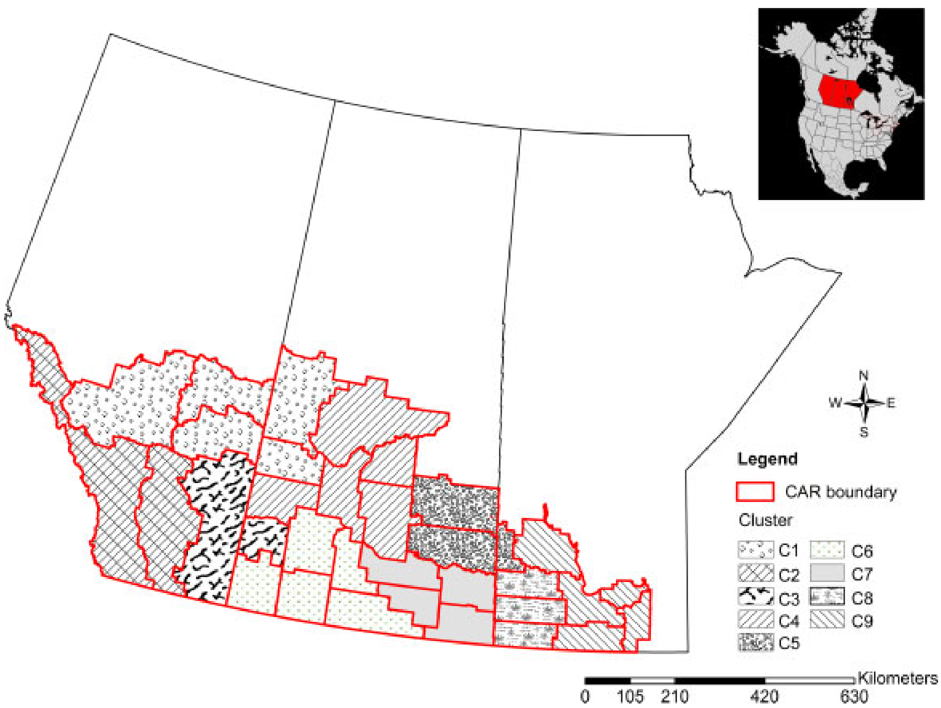

Figure 1. Spatial distribution of the 34 CARs across the prairies. Nine clusters (C1–C9) were composed of neighbouring CARs as determined by Ward's (1963) minimum-variance cluster analysis. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

The Canadian prairies cross the southern parts of the provinces of Alberta, Saskatchewan, and Manitoba. Dominated by a semi-arid climate, the precipitation in this region is highly variable and unevenly distributed. Growing season (May to August) precipitation averages about 200 mm, lower than the crop water demand (approximately 300 mm) (Ash *et al*., 1992). Winter precipitation in the form of snowfall is important, accounting for approximately one third of annual precipitation and producing 80% or more of annual local surface runoff (Pomeroy and Goodison, 1997). The prairie landscape is dominated by plains and gently rolling terrain. The absence of any significant topographic barrier in the vast north–south corridor is responsible for the great variety of weather (Hare and Thomas, 1919).

2.2. Data

Spring wheat yield data (tha^{-1}) for each CAR, available for 28 years (1976–2003), were obtained from Statistics Canada. Spring wheat was selected in this study because its acreage is the highest compared with other crops in the prairies and it is growing extensively in all CARs. Also, spring wheat is widely used in the literature to represent agricultural drought on the prairies (e.g. Kumar and Panu, 1997; Quiring and Papakyriakou, 2003).

Three widely used drought indices, the Standardized Precipitation Index (SPI; McKee *et al*., 1993) at three time steps (1-, 3-, and 6-month), the Palmer Drought Severity Index (PDSI; Palmer, 1965), and the Palmer Moisture Anomaly Index (Z-index; Palmer, 1965) were selected for use in this study. Their individual effectiveness in characterizing agricultural drought on the prairies has been widely studied (e.g. Akinremi *et al*., 1996; Quiring and Papakyriakou, 2003; Wheaton *et al*., 2008). The SPI's calculation requires at least 30 years of monthly precipitation data for the region. Besides precipitation, the PDSI and Z-index's calculations require at least 30 years' monthly maximum and minimum temperature data, plus the available holding capacity (AWHC) of the soil and the longitude and latitude of the site for which it is being calculated.

Considering weather stations are unevenly distributed in some CARs and this affects the reliability of CAR averaged drought indices, daily maximum and minimum temperature and precipitation data were obtained from the daily 10 km gridded climate data set for Canada (1961–2003; AAFC, 2008a). Grids were interpolated from daily Environment Canada climate station observations using a thin plate smoothing spline surface fitting method implemented by ANUSPLIN V4.3. This method has been shown to perform well when interpolating noisy climate data across complex terrain in comparison with other interpolation techniques (Hutchinson and Gessler, 1994; McKenney *et al*., 2006). According to the metadata of this specific gridded climate data set, on average, ANUSPLIN tends to overestimate extreme minimum temperature by around 0.6 °C, underestimate extreme maximum temperature by around 0.25 °C, and underestimate high precipitation extremes by around 2 mm.

The SPI was calculated using FORTRAN 90/95 code provided by the National Agroclimate Information Service (NAIS) of Agriculture and Agri-Food Canada (AAFC). The AWHC value of the soils was defined via the Soil Landscapes of Canada (SLC) Version 3.1.1 (AAFC, 2008b).

The PDSI and Z-index were calculated using the National Drought Model employed by the NAIS Drought Watch program. The PDSI calculated from Palmer's original method has considerable limitations, including the use of two simplified soil layers in the water balance computations that may not be accurately representative of a location, the use of empirical constants for the climatic characteristic and the duration factors, limiting the spatial comparability of the index, and the estimation of potential evapotranspiration with the Thornthwaite method (Thornthwaite, 1948) which yields less realistic estimates than Priestley and Taylor's (1972) method (Alley, 1984; Guttman *et al*., 1992). The most recent version of PDSI model developed by Wells *et al*. (2004) improved the spatial comparability of PDSI values by replacing the empirical constants in the index with dynamically calculated values. The national drought model version used in this study is based on a six-layer structure that is more accurate in tracking the movement of soil moisture than Palmer's two-layer model. It also replaces the regional correction factor of 17.67 employed in Palmer's original method to 14.2, established by Akinremi *et al*. (1996), to better simulate soil moisture variations in Canada. Moreover, the potential evapotranspiration is calculated using the Priestley and Taylor equation which has better physical appeal than the Thornthwaite's method. These alterations overcome many of the limitations in the Palmer drought indices, providing a more accurate measurement of moisture conditions of the study area.

Five drought indices were calculated for the entire period from 1961 to 2003 to characterize long-term conditions, but due to the availability constraints over the entire study area, only the data from 1976 to 2003 were used for further analysis. Gridded drought indices were then aggregated to CAR averages, by taking the average value of all grid cells within each CAR.

2.3. Methods

2.3.1. *Crop yield data detrending and standardization*

Owing to advances in agricultural technology, such as greater rates and frequency of fertilizer application, the use of new crop varieties, improved weed control, and better tillage practices, agricultural areas are generally experiencing an upward trend in spring wheat yields (Qian *et al*., 2009). To eliminate bias due to non-climatic factors, the trend was removed using linear regression when calculating yield variability (e.g. Hill *et al*., 1980; Wu *et al*., 2004). To compare yield variability from CARs with different mean values and standard deviations, the yield residuals were standardized for each CAR using the Z-score transformation quantifying the original score in

Table I. Canadian prairies drought intensity classification.

Standardized yield residuals	Drought category	Cumulative frequency (%)
> −0.20	Non-drought	>35
> −0.69 to −0.20	Mild	>20 to 35
> −1.24 to −0.69	Moderate	>10 to 20
> −1.84 to −1.24	Severe	>5 to 10
≤ −1.84	Extreme	≤5

terms of the number of standard deviations that the score is from the mean of the distribution.

2.3.2. *Agricultural drought intensity classification*

Drought intensity is one of the essential components for representing a comprehensive picture of drought for a region (Sivakumar and Whilhite, 2002). In this study, a five-level drought intensity classification – non-drought, mild, moderate, severe, and extreme – was utilized (Table I). As we focused on assessing dry conditions, the near normal and wet conditions were grouped into one class termed 'non-drought.' The chance of occurrence for each level of drought was defined according to the commonly used cumulative frequency of different drought intensities, such as the percentile categories employed by the U.S. Drought Monitor (Svoboda *et al*., 2002) and the SPI classification (McKee *et al*., 1993). For each drought category, the corresponding yield threshold with the specific chance of occurrence was calculated from an empirical cumulative distribution function developed from the standardized spring wheat yields of the 34 CARs during the 28 years. According to this classification, a year is identified as a drought year with a specific intensity when the corresponding yield is lower than the historical mean by 0.20 standard deviations.

2.3.3. *Creation of spatial scale for agricultural drought risk assessment*

Widely used spatial scales for drought evaluation include climate divisions (Svoboda *et al*., 2002; Steinemann and Cavalcanti, 2006), crop districts (Quiring and Papakyriakou, 2003), and political jurisdictions (e.g. countries, provinces) (Wu *et al*., 2004). For the prairies, evaluating droughts by province or territory is not appropriate, because administrative boundaries do not accurately reflect the physical features and climate of the region. To increase the sample size for model development, neighbouring CARs were grouped together based on the similarity of yield variations throughout the study period.

Ward's (1963) minimum-variance hierarchical cluster analysis was applied to determine which CARs were most similar by maximizing the proportion of variation in standardized yields explained by a particular clustering of the CARs. Different clustering solutions were examined and a nine-clustering solution was selected, because it kept as much similarity of the CARs within each cluster as possible, and the number of CARs was relatively equal within each cluster. As shown in Figure 1, the clusters appeared to be primarily controlled by geographic location, conforming to the generalization that geographically close regions usually experience similar weather and crop response (Wu *et al*., 2004; Williams *et al*., 2008). These clusters, referred to as C1, C2 ... C9, were used as the spatial units for model development.

2.3.4. *Identification of outliers*

Both drought and flood can lead to crop water stress and thus reduce crop yield. As the drought indices chosen in this study do not reflect the effects of flooding, the flood-induced yield reductions, which have an inconsistent influence on the model development, needed to be removed from the analysis. The criterion for identifying the influential data was the consistency between the standardized yield residuals and the drought indices values before and during the growing season (c.f. Wu *et al*., 2004). For each CAR, in cases where the drought indices were constantly over 2 (i.e. moderately wet) for at least 2 months, but the standardized yield residual was lower than −0.69 (i.e. associated with moderate drought), the data of that year were omitted from model development. For each CAR, the number of flood years varies from 0 to 3 during the study period.

For example, 1997 was a severe flood year in Manitoba. The Red River flooding severely damaged farm buildings, equipment, and delayed spring planting in rural areas (Environment Canada, 2008), and as a result the yield was very poor in this year (e.g. the standardized yield residual for CAR 4609 was −1.04). However, the drought indices (expressed in terms of yield) were extremely high in affected CARs, which would predict a high-yield year. Therefore, the data from 1997 were identified as outliers for most of the CARs in Manitoba and were not used to develop the model. Replacing our approach with a model of flood impacts on crop yield would be an appropriate method to fill in the resulting gaps.

2.3.5. *Multi-index drought model development*

A multi-index drought (MID) model was developed to predict agricultural drought at six stages: pre-planting, and at the beginning of each month during the growing season (defined in the model as May 1st to September 1st). Before the growing season, the MID model assesses drought risk at the beginning of April by evaluating the drought indices of the recharge period (from the previous September to current March for a given year). During the growing season, the model is updated at the beginning of each month by assessing the drought indices from the recharge period to the preceding month. The last stage of the model is updated at the beginning of September, using the drought indices from August.

Principal Component Analysis (PCA) and multiple linear regression were used to establish a predictable relationship between drought indices and the standardized spring wheat yield residuals. A nested loop procedure

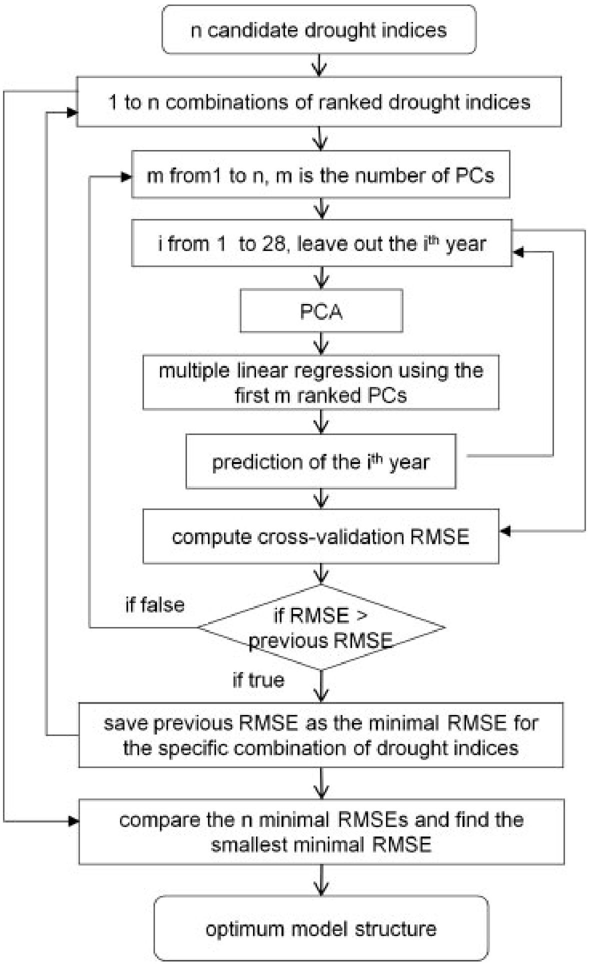

Figure 2. Flowchart of the nested loop approach for MID model construction, where n is the number of candidate drought indices at each stage of modelling; m is the number of ranked resulting principal components (ranked by their correlations between standardized yield residuals) for the multiple regression analysis. In each iteration of cross-validation (CV), a different year, i, was withheld for validation until all 28 years were used for validation. For each modelling stage, the optimum model structure was obtained by minimizing the CV RMSE in the nested loop approach.

was employed to obtain the optimum model structure (Figure 2). The outer loop selects a robust combination of predictors. For each specific combination of drought indices, the inner loop controls the selection of which and how many principal components (PCs) should be used for regression analysis. For a total of 28 years, a leave-one-out cross-validation (CV) (by year) approach was used to evaluate the model performance. For each stage of modelling, the optimum model structure (i.e. the robust combination of drought indices and the best resulting PCs for the regression) was determined by the minimized CV RMSE (i.e. root mean square error from the cross validation).

This approach evaluates different combinations of candidate drought indices at each stage of modelling, which was necessary because as the crop develops, the impact of previous months' weather conditions on yield becomes less important and may only introduce noise. To save computing resources, the evaluations were not performed on all possible combinations of candidates (i.e. different combinations from drought indices with poor correlations between yield residuals were ignored). The first evaluated combination was a single drought index with the highest correlation between yield residuals. The subsequent combinations were created by adding the next highest ranked drought index at each loop step, until all candidates were included.

However, the specific combinations of drought indices could not be used directly in a multiple regression, because of the inherent multicollinearity among drought indices. Multicollinearity refers to a situation in which two or more explanatory variables in a multiple regression model are highly correlated, with a large portion of shared variance and low levels of unique variance (Coolidge, 2000). PCA was employed to extract the smallest number of uncorrelated components (i.e. PCs) that account for most of the variation in the original multivariate data (Rogerson, 2001). Although the first several PCs account for most of the total variance, there is no guarantee that they are the best predictors, and low variance components can have significant correlation with a dependant variable (Jollife, 1982; Hadi and Ling, 1998).

To select the best PCs for multiple regression analysis, another optimization loop (i.e. the inner loop of the nested approach) was employed. Similar to the drought index selection process, the PCs were ranked by comparing their correlations between the standardized yield residuals. The evaluation began from the highest-ranking PC, and then iteratively added the next highest-ranking PC, until the CV RMSE started to increase. In this study, the minimum CV RMSE was achieved with no more than four highest-ranking PCs (one or two PCs for the majority of cases). For each stage of modelling, a set of minimum CV RMSEs from different combinations of drought indices were compared, and the optimum model structure was defined as the specific combination of drought indices and specific number of resulting ranked PCs that obtained the lowest minimum CV RMSE in the set.

Three model evaluation statistics were employed to evaluate the prediction accuracy, including the coefficient of determination (R^2), the RMSE, and the prediction accuracy associated with each drought category. A leave-one-out cross-validation was employed again to evaluate the prediction accuracy of the model (c.f. Qian *et al.*, 2009). For each modelling stage, using the final-selected predictors and the best PCs, the predicted results for each year were obtained from the model calibrated with the remaining 27 years of data.

3. Results and discussion

3.1. Results of cross-validation of the MID model

The R^2 and standardized yield residual RMSE results for the 54 MID models (9 clusters $\times$ 6 stages) are plotted by cluster (Figure 3). The average rate of drought category prediction accuracy for each stage of modelling is summarized in confusion matrices (Table II).

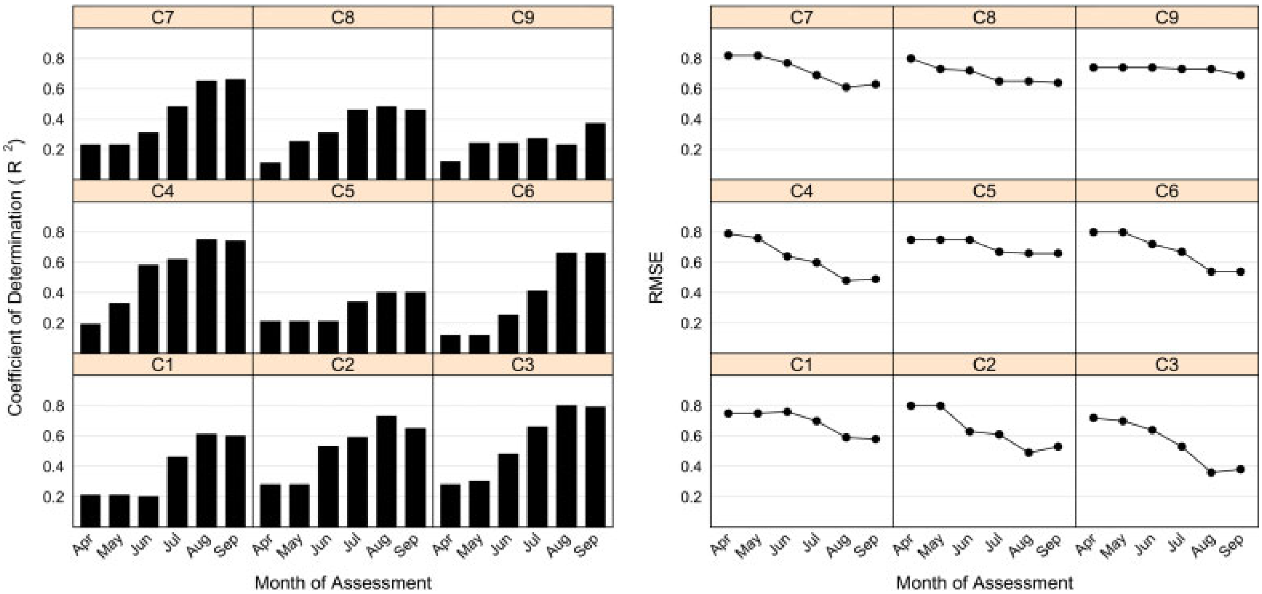

Figure 3. Plots of the R^2 (left) and RMSE (right) between the observed and the predicted standardized spring wheat yield residuals by cluster. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Unsurprisingly, due to the uncertainties of (future) growing season weather conditions, the model performance at the pre-planting stage was relatively poor compared to that of the later stages. Only 20% of mild and 11% of moderate droughts were precisely predicted, 42% of severe droughts were detected as mild or moderate, and 44% of extreme droughts were forecasted to be mild to severe at this stage. This indicates that while growing season's weather is critical to evolving drought conditions, the recharge period weather conditions also have an important impact on drought occurrence and persistence. This is understandable, because growing season precipitation is usually not sufficient to meet crop demand for most regions on the prairies. If spring soil moisture levels are drier than normal, timely above-normal precipitation is required during the growing season to make up the deficit (Sutton, 2003). However, the chance of this occurring in a growing season is low on the prairies. Therefore, drought indices derived from recharge period are valuable for agricultural drought risk assessment, providing a warning of possible drought progression as early as April.

When compared with Stage 1 and with regards to R^2, the model performance at Stages 2 and 3 increased for the most regions, but remained stable for C1 and C5. This indicates that the contributions of April and May weather conditions on the yields varies spatially. For some regions, the contributions are minor. The assessment accuracy was further improved at the fourth stage and reached its highest at the fifth stage (average $R^2 = 0.60$). It is not surprising that the strongest correlation between the drought indices and spring wheat yield residuals was in June (Stage 4) and July (Stage 5), because spring wheat yield is largely determined by moisture stress during the heading and soft dough stages, which usually occur during the second half of June through July (Arora *et al.*, 1987; Quiring and Papakyriakou, 2003). Crops at these stages are vulnerable to drought, and even a moderate drought may reduce the yield greatly. Therefore, June and July are the most important months for determining the risk of agricultural drought. Converting the standardized yield residuals of Stage 5 to yield residuals, the mean yield residual RMSE was 0.238 tha^{-1}, ranging from 0.162 to 0.387 tha^{-1}.

Mavromatis (2007) evaluated the SPI and three variations of the PDSI (the original PDSI, a self-calibrated version, and a modified scheme employing Priestley–Taylor's approach to compute potential evapotranspiration instead of Thornthwaite's method) and their respective Z-index for accessing rain-fed common wheat and durum wheat yield in two pilot regions in north and central Greece. The model performance statistics showed that the drought indices based on Palmer's scheme are most suitable for predicting yields. The self-calibrated PDSI ranked the highest (RMSE = 0.105 tha^{-1}) for predicting durum wheat residuals and the original PDSI ranked the highest (RMSE = 0.149 tha^{-1}) for predicting common wheat residuals. Greece has a sharply seasonal Mediterranean climate, and the summer is extremely hot and rainless. Compared to the Canadian prairies, moisture stress is a more significant factor in limiting yield in Greece, and thus the prediction accuracy of Palmer's drought indices was higher. Quiring and Papakyriakou (2003) performed an evaluation of four drought indices (SPI, PDSI, Z-index and NDI (NOAA Drought Index) for predicting spring wheat yield on the Canadian prairies based on the sum of the index's June and July values. The model evaluation indicated that the Z-index was the most appropriate index for predicting yield departures when there is significant moisture stress, with a mean RMSE value of 0.256 tha^{-1} for all 43 crop districts. The multi-index approach was able to achieve lower mean RMSE, however, of 0.238 tha^{-1}.

Table II. The average prediction accuracy rate (%) for each drought category across all clusters.

Actual (%)		Predicted (%) Non-drought	Mild	Moderate	Severe	Extreme
Stage 1	Non-drought	75	20	4	1	0
	Mild	73	20	7	0	0
	Moderate	53	36	11	0	0
	Severe	58	38	4	0	0
	Extreme	56	33	9	2	0
Stage 2	Non-drought	74	21	4	1	0
	Mild	73	19	8	0	0
	Moderate	64	24	11	1	0
	Severe	60	35	5	0	0
	Extreme	58	33	5	4	0
Stage 3	Non-drought	77	16	6	1	0
	Mild	67	22	11	0	0
	Moderate	54	32	12	2	0
	Severe	42	35	20	3	0
	Extreme	42	42	9	7	0
Stage 4	Non-drought	79	16	5	0	0
	Mild	57	30	11	2	0
	Moderate	46	21	26	7	0
	Severe	27	28	30	15	0
	Extreme	35	19	30	16	0
Stage 5	Non-drought	82	15	3	0	0
	Mild	54	35	9	2	0
	Moderate	25	36	29	10	0
	Severe	20	20	35	20	5
	Extreme	14	31	23	23	9
Stage 6	Non-drought	81	15	3	1	0
	Mild	58	32	7	3	0
	Moderate	25	32	37	6	0
	Severe	11	18	45	23	3
	Extreme	4	23	40	26	7

August weather conditions seemed to have little contribution to the prediction accuracy of Stage 6, and even had a negative influence for some regions (e.g. C2). This suggests that late summer weather plays only a minor role in agricultural drought prediction, or may even mislead the assessment. This is consistent with the findings of Whitmore (2000), who pointed out that drought has little further detrimental effect on the wheat from the hard dough stage up to ripening. At this point, spring wheat is near maturity and does not respond to water stress as much as during the previous stages. A slightly drier than normal August ensures that harvest can take place without difficulty or significant loss of yield (Whitmore, 2000).

3.2. Spatial variability of model performance

The results also revealed great spatial variation across the prairies in the model performance (Figure 4), generally paralleling the pattern of growing season precipitation. Figure 5 shows the mean and the coefficient of variation (standard deviation/mean) of growing season precipitation for the period of 1976 to 2003. The south, southwest, and central prairies that receive inadequate amounts of growing season precipitation and have a large year-to-year variability in precipitation regime tend to have stronger correlations between the drought indices and the spring wheat yield. Conversely, the correlation becomes weak in the eastern prairies with abundant and reliable precipitation spread over the growing season.

In addition, southern prairies tend to have more fertile soils and higher temperature (Acton *et al*., 1998; Luo and Zhou, 2006). These factors, which were not considered in the model, do not normally limit crop growth in the south, but may have increased influence on the crop production of more northerly prairies. All these factors collectively influence the correlation between spring wheat yield and drought indices, and should be jointly considered when interpreting their relationships to model performance.

3.3. Comparison of multi-index and single drought index

To further explore the advantages of employing multiple drought indices versus a single drought index in this

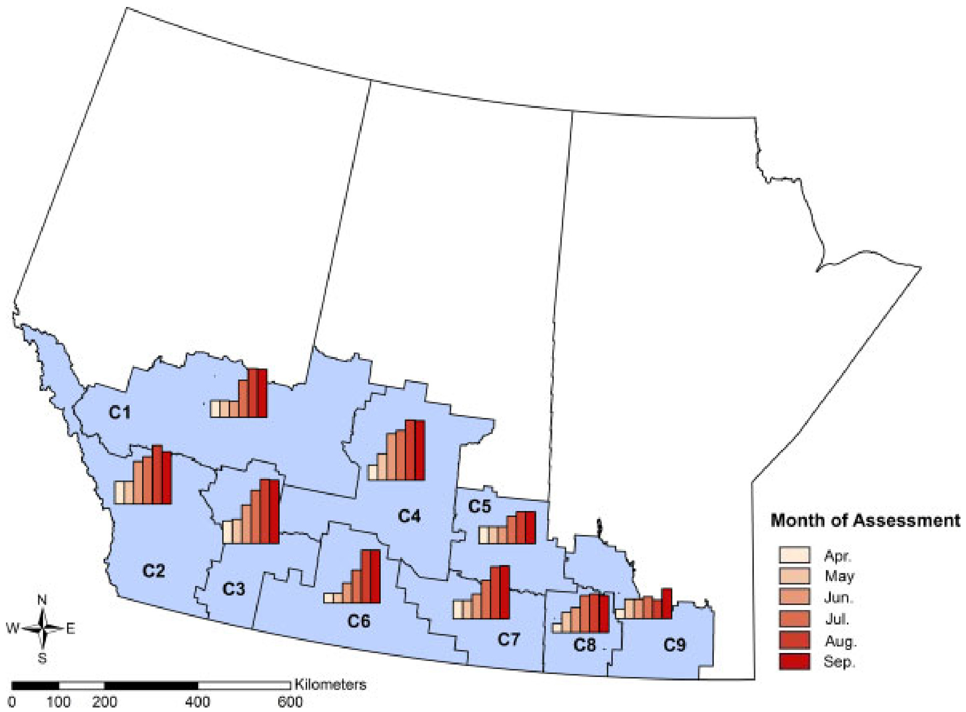

Figure 4. Mapping the R^2 for the relationship between multiple drought indices and the standardized spring wheat yield residuals. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

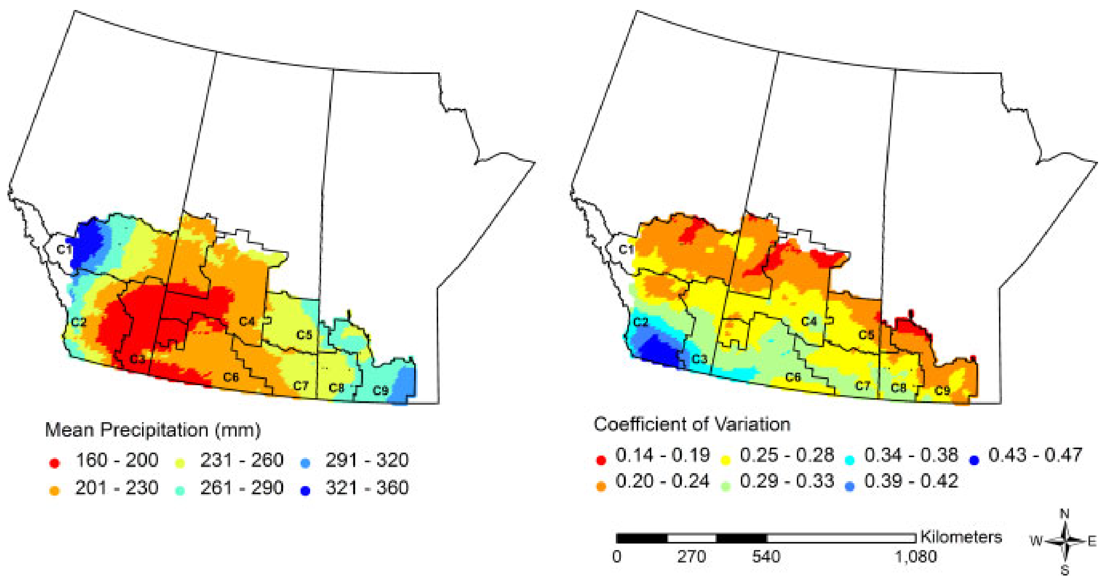

Figure 5. Mapping the mean (left) and coefficient of variation (right) of growing season (May 1st to September 1st) precipitation for the prairies (1976–2003). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

study, the model was run using each individual drought index separately. As can be seen from Figure 6, in general the CV RMSEs of multi-index were lower than (or occasionally equal to) any individual drought index for all regions. Employing multiple drought indices is especially beneficial for more arid areas, such as C3, where there is a significant difference between the CV RMSE from using multiple and single drought index. For more humid areas that consistently receive large amounts of precipitation during the growing season (i.e. soil moisture is no longer an important yield-limiting factor), such as C9, the performance of any single drought index is relatively weak and thus the advantages of combing multiple drought indices are less significant.

The comparison confirmed that the effectiveness of any single drought index is temporally and spatially dependent. For example, the PDSI was more suitable for use in cluster 4 than in cluster 1. Even within specific regions, such as C3, the PDSI was less valuable at early stages than at latter stages. This is in accordance with the findings of Mavromatis (2007), who pointed out that choosing the best suited index is particularly difficult because

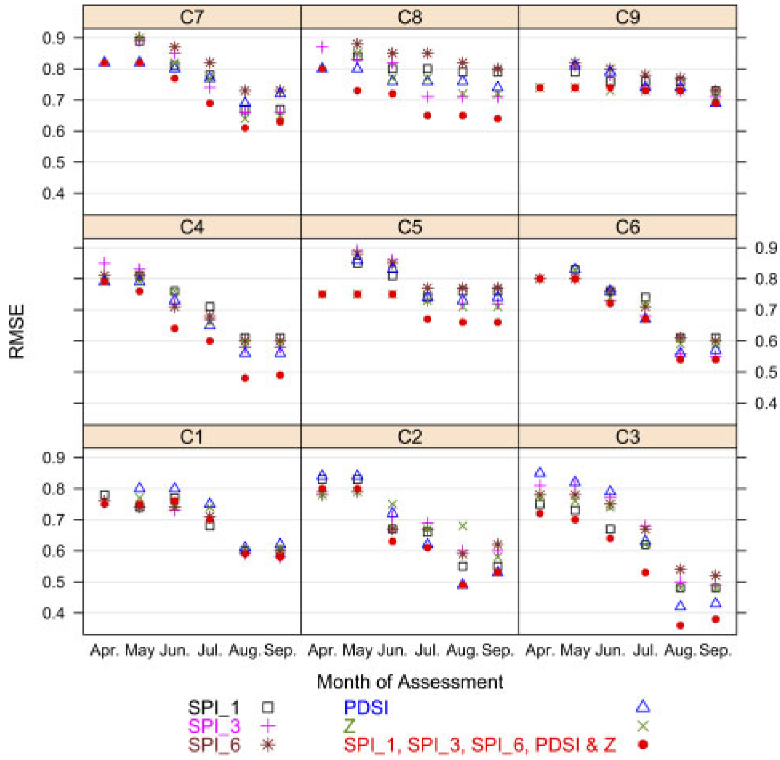

Figure 6. Comparison of the model CV RMSE from each individual drought index (i.e. SPI_1, SPI_3, SPI_6, PDSI and Z-index, respectively) and the multiple drought indices. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

the answer will vary depending on the crop's sensitivity to moisture storage and the characteristics of the study area (e.g. soil properties/variability, climate regime, etc.). Therefore, employing multi-index approach in drought risk management is beneficial in overcoming some of the deficiencies in any individual index.

3.4. Limitations of the MID model

The MID model is weak at accurately predicting large negative yield departures, as severe and extreme droughts were routinely underpredicted (Table II). For example, none of the extreme droughts was precisely predicted until Stage 5, with an accuracy rate of 9%, and 77% of them were underestimated as mild to severe droughts. One possible reason for this is the insufficient number of observations with extremely low yield. For each region, there are not more than 3 years with extreme low yield within the 28 study years. It is also likely that the response of yield to dry conditions may not be linear when the soil moisture drops below a certain threshold. A nonlinear model may be more appropriate to estimate yields under very dry conditions (e.g. Quiring and Papakyriakou, 2003). For regions with good density and relatively evenly distributed weather stations, the use of in-situ station data would further improve the prediction accuracy of severe and extreme droughts.

The second weakness is that the model's prediction accuracy did not always show a stable increase from stage to stage during the growing season. A certain level of drought that had been correctly predicted at earlier stages was likely to be mis-predicted as other categories at later stages. This may be due to the model's monthly scale of assessment, which limits the opportunity to predict droughts that occur for shorter intervals or to associate the water stress to critical growth stages that are less than a month. Besides total precipitation, the distribution of the precipitation over the growing season is critical to crop yield. In some cases, several large precipitation events will skew the monthly precipitation totals, and the empirical nature (specifically, the temporal resolution of the relationships) of the model may miss cases of inadequate precipitation in certain critical, water-sensitive periods (e.g. at the end of June and the beginning of July). Therefore, above average but poorly distributed growing season precipitation can also lead to poor yield if the timing prevents proper crop development or agricultural practices. Conversely, even if the total growing season precipitation complies with a numerical definition of drought, it could possibly be so well-distributed in terms of a crop's pattern of water demands, that it provides an adequate or even superior crop yield (Whitmore, 2000). It is likely that the model performance would be further improved if drought conditions were assessed weekly or biweekly.

In addition, some poor predictions may be attributed to factors other than drought, such as pests, disease, weeds, and other weather-related damages (e.g. hail, wind, and frost). Furthermore, the predictions could be affected by the quality of the weather and observed yield data. The observed yield data have errors that are particularly difficult to quantify due to the constraints of privacy protection.

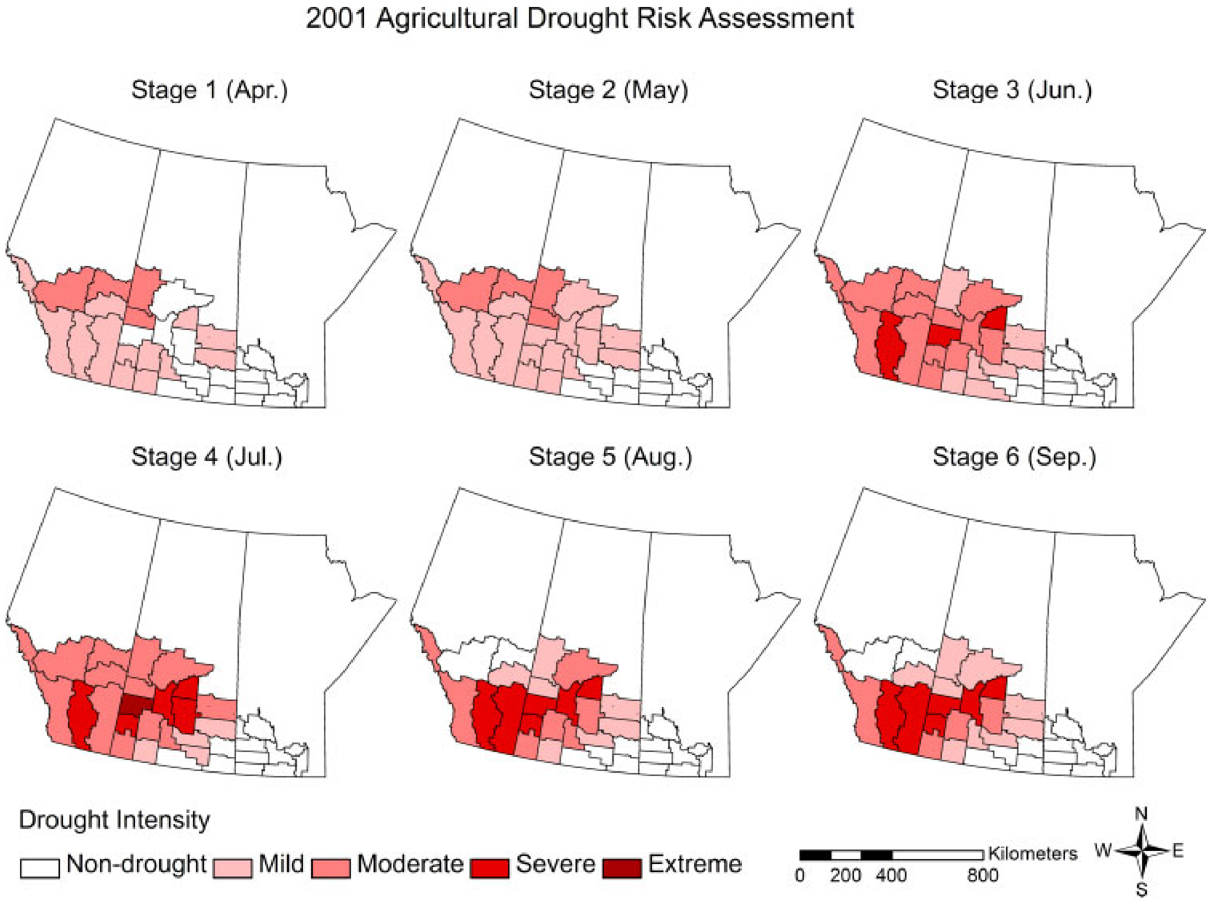

Figure 7. Mapping of the MID modelling results of 2001. The results were obtained at the beginning of each month from April to September. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

3.5. MID model application

To provide a better visualization of the assessment results, the MID model was applied to 2001, which was identified as the most severe drought on record in parts of the prairies (Bonsal and Wheaton, 2005).

The spatial patterns and temporal behaviour of the predictions are shown in Figure 7. Alberta and part of Saskatchewan were identified as being under mild to moderate drought stress at early stages. As the crop developed, the increasingly severe drought risk spread to most of the regions in Saskatchewan. Compared to the actual drought conditions of 2001, which were determined by the observed standardized spring wheat yield residuals (Figure 8), the overall assessment generally resembled the dry conditions of 2001.

4. Conclusions

This article aims to present an operational model framework that combines the strengths of various drought indices to provide a more comprehensive assessment of agricultural drought conditions in the Canadian prairies. The MID model combines the strengths of various drought indices derived from preceding and current months as they become available to better match the conditions (both spatially and temporally) in which they work well, providing a more reliable and comprehensive assessment of drought conditions. The results showed that (1) the prediction accuracy of the MID model is better than (or occasionally equal to) using any single drought index for all modelling stages, (2) drought indices derived from recharge period are useful for early drought risk detection, (3) the prediction accuracy improved as the growing season progressed, with the most accurate assessments at the beginning of August, and (4) a multi-index approach is best suited for the more arid locations in the southern prairies, which tend to have a more variable precipitation regime.

Further improvements in the approach presented here are likely by generating drought indices on shorter time scales (weekly or biweekly) to assess short-term dry spells during critical crop phenological stages. A non-linear model may be more appropriate to estimate crop yield under very dry conditions. Additional data that measure some of the other factors affecting crop yield, such as pests, diseases, and weather-related damages would be valuable to further increase model performance. It would also be useful to extend this study to cover other crops that are important to the prairies, such as canola and barley.

The drought and responding crop yield patterns vary through time with climate and cropping pattern changes. If the MID model is used on an operational basis, the clusters need to be updated periodically as more years of

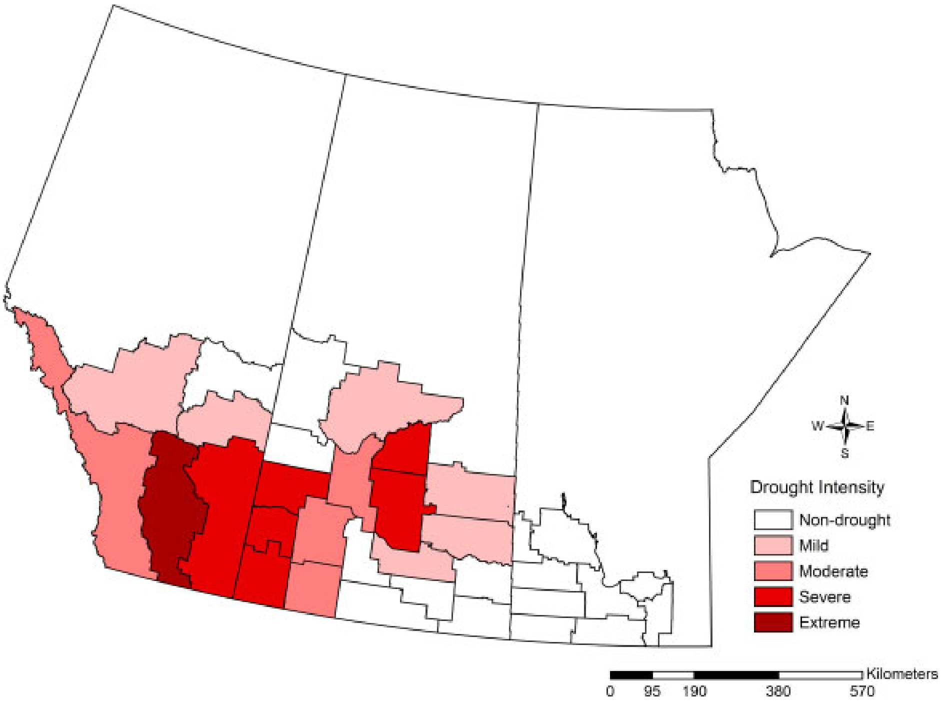

Figure 8. Spatial distribution of agricultural drought intensity of 2001 as determined by the observed standardized spring wheat yield residuals. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

yield data become available. The model is not suitable for drought-tolerant crops, because for them soil moisture is not as important as a yield-limiting factor. It is also noted that the MID model is not suitable for use in flood-prone regions, because drought indices do not reflect flood-induced yield reductions. However, a combination of the MID with models of flood impacts would be an appropriate subsequent improvement.

Besides climate-based drought indices that can be derived from relatively easily available weather data, complementary data from remote sensing, radar, and other technologies such as biophysical models should also be explored. Compared to climate-based drought indices, satellite-derived data have high availability over large areas, which is particularly important in data scarce regions (Johnson *et al*., 1993). However, optical remote sensing techniques are of limited value to assess recharge period weather conditions, drought stress before seeding and near harvest. Radar-based soil moisture estimates offer promise for initial assessment of available water prior to planting, but Boisvert *et al*. (1997) points out that these estimates are at shallower depths than spring wheat rooting depth. Near harvest, remote sensing vegetation indices approach a saturation level during the last stage of crop development, making this method less effective than desired (Qian *et al*., 2009). Process-based models are often limited by data needs or problems with parameterizing the models with confidence over large spatial extents, such as those presented in this study. Since all of these methods, including the MID model presented in this article, show weaknesses in particular situations, it will continue to be important to evaluate and combine a variety of approaches to predict and monitor agricultural drought over large areas such as the Canadian prairies.

Acknowledgements

This study was financially supported by the Government Related Initiatives Program – 'Integrating remote sensing data into selected models to enhance operational decision support for crops, drought and agricultural water management' of Agriculture and Agri-Food Canada (AAFC). The authors would like to thank Richard Warren (AAFC) for providing the computer program to generate drought indices, and Frédéric Bédard (Statistics Canada) for providing the spring wheat yield data and clarified related issues. Thanks also to Budong Qian, Andrew Bootsma, Nathaniel Newlands, Xiaoyuan Geng, Allan Howard, Aston Chipanshi, and Zhirong Yang (AAFC) for their valuable suggestions and comments.

References

AAFC (Agriculture and Agri-Food Canada). 2008a. *Gridded Climate Data*, http://www4.agr.gc.ca/AAFC-AAC/display-afficher.do?id=1227620138144&lang =eng

AAFC. 2008b. *Soil Landscapes of Canada Version 3.1.1*, http://sis2.agr.gc.ca/cansis

Acton DF, Padbury GA, Stushnoff CT. 1998. *Prairies*, http://esask.uregina.ca/entry/prairies.html

Akinremi OO, McGinn SM, Barr AG. 1996. Evaluation of the Palmer drought index on the Canadian prairies. *Journal of Climate* **9**: 897–905.

Alley WM. 1984. The Palmer drought severity index: limitations and assumptions. *Journal of Applied Meteorology and Climatology* **23**: 1100–1109.

Arora VK, Prihar SS, Gajri PR. 1987. Synthesis of a simplified water use simulation model for predicting wheat yields. *Water Resources Research* **23**(5): 903–910.

Ash GHB, Shaykewich CF, Raddatz RL. 1992. Moisture risk assessment for spring wheat on the eastern prairies: a water-use simulation model. *Climatological Bulletin* **26**(2): 65–78.

Boisvert JB, Gwyn QHJ, Chanzy A, Major DJ, Brisco B, Brown RJ. 1997. Effect of surface soil moisture gradients on modelling radar backscattering from bare fields. *International Journal of Remote Sensing* **18**(1): 153–170.

Bonsal BR, Wheaton EE. 2005. Atmospheric circulation comparisons between the 2001 and 2002 and the 1961 and 1988 Canadian prairie droughts. *Atmosphere-Ocean* **43**(2): 163–172.

Bordi I, Sutera A. 2007. Drought monitoring and forecasting at large scale. In *Methods and Tools for Drought Analysis and Management*, Rossi G, Vega T, Bonaccorso B (eds). Springer Netherlands: New York; 3–27.

Brown JF, Wardlow BD, Tadesse T, Hayes MJ, Reed BC. 2008. The vegetation drought response index (VegDRI): a new integrated approach for monitoring drought stress in vegetation. *GIScience and Remote Sensing* **45**: 16–46.

Cancelliere A, Di Mauro G, Bonaccorso B, Rossi G. 2007. Drought forecasting using the standardized precipitation index. *Water Resources Management* **21**(5): 801–819.

Coolidge FL. 2000. *Statistics: A Gentle Introduction*. Sage: London.

Environment Canada. 2008. *Flooding events in Canada – Prairie Provinces.* http://www.ec.gc.ca/eau-water/default.asp?lang=En&n=E0399791-1

Guttman NB, Wallis JR, Hosking JRM. 1992. Spatial comparability of the Palmer drought severity index. *Water Resources Bulletin* **28**: 1111–1119.

Hadi AS, Ling RF. 1998. Some cautionary notes on the use of principal components regression. *The American Statistician* **52**: 15–19.

Hare FK, Thomas MK. 1919. *Climate Canada*. John Wiley and Sons Canada Limited: Toronto.

Heim RR. 2002. A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society* **83**(8): 1149–1165.

Hill JD, Strommen ND, Sakamoto CM, Leduc SK. 1980. LACIE – an application of meteorology for United States and foreign wheat assessment. *Journal of Applied Meteorology* **19**: 22–34.

Hutchinson MF, Gessler PE. 1994. Splines – more than just a smooth interpolator. *Geoderma* **62**: 45–67.

Johnson GE, Achutuni VR, Thiruvengadachari S, Kogan FN. 1993. The role of NOAA satellite data in drought early warning and monitoring: selected case studies. In *Drought Assessment, Management, and Planning: Theory and Case Studies*, Wilhite DA (ed). Kluwer Academic Publishers: Boston, 31–48.

Jollife IT. 1982. A note on the use of principal components in regression. *Applied Statistics* **31**(3): 300–303.

Kumar V, Panu U. 1997. Predictive assessment of severity of agricultural droughts based on agro-climatic factors. *Journal of the American Water Resources Association* **33**(6): 1255–1264.

Luo Y, Zhou X. 2006. Temperature and spatial variations in soil respiration. In *Soil Respiration and the Environment*, Luo Y, Zhou X (eds). Academic Press: New York, 107–132.

Mavromatis T. 2007. Drought index evaluation for assessing future wheat production in Greece. *International Journal of Climatology* **27**(7): 911–924.

McKee TB, Doesken NJ, Kleist J. 1993. The relationship of drought frequency and duration to time scales. In Eighth Conference on Applied Climatology, Anaheim, California, 179–184.

McKenney DW, Pedlar JH, Papadopol P, Hutchinson MF. 2006. The development of 1901–2000 historical monthly climate models for Canada and the United States. *Agricultural and Forest Meteorology* **138**: 69–81.

Morgan R. 1985. The development and application of a drought early warning system in Botswana. *Disasters* **9**(1): 44–50.

Morid S, Smakhtin V, Moghaddasi M. 2006. Comparison of seven meteorological indices for drought monitoring in Iran. *International Journal of Climatology* **26**: 971–985.

Nkemdirim L, Weber L. 1999. Comparison between the droughts of the 1930s and the 1980s in the southern prairies of Canada. *Journal of Climate* **12**(8): 2434–2450.

Palmer WC. 1965. Meteorological drought. Research Paper No. 45, US Weather Bureau, Washington, DC.

Pomeroy JW, Goodison BE. 1997. Winter and snow. In *The Surface Climates of Canada*, Bailey WG, Oke TR, Rouse WR (eds). McGill-Queen's University Press: Montreal, 68–100.

Priestley CHB, Taylor RJ. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review* **100**: 81–92.

Qian B, De Jong R, Warren R, Chipanshi A, Hill H. 2009. Statistical spring wheat yield forecasting for the Canadian prairies provinces. *Agricultural and Forecast Meteorology* **149**(6–7): 1022–1031.

Quiring SM. 2009. Monitoring drought: an evaluation of meteorological drought indices. *Geography Compass* **3**(1): 64–88.

Quiring SM, Papakyriakou TN. 2003. An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and Forest Meteorology* **118**(1–2): 49–62.

Rogerson PA. 2001. Data reduction: factor analysis and cluster analysis. In *Statistical Methods for Geography*, Rogerson PA (ed). Sage: London, 192–197.

Sinha CP, Chaube UC, Saxena RP. 1992. A critical analysis of drought indices for different regions. In *Proceedings World Congress on Natural Hazards*, New Delhi, 10–14.

Sivakumar MKV, Wilhite DA. 2002. Drought preparedness and drought management. In *International Conference on Drought Mitigation and Prevention of Land Desertification*, Bled, Slovenia.

Statistics Canada. 2003. *Geographic Units*. http://geodepot.statcan.ca/Diss/Reference/COGG/Index_e.cfm [2009].

Steinemann AC, Cavalcanti LFN. 2006. Developing multiple indicators and triggers for drought plans. *Journal of Water Resources Planning and Management* **132**(3): 164–174.

Steinemann AC, Hayes MJ, Cavalcanti LFN. 2005. Drought indicators and triggers. In *Drought and Water Crises: Science, Technology, and Management Issues*, Wilhite DA (ed). CRC Press: Boca Raton, 71–92.

Sutton D. 2003. *Agroclimatic Atlas of Alberta: Soil Moisture Conditions in Alberta*. http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/sag6302 [2009].

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

Thornthwaite CW. 1948. An approach toward a rational classification of climate. *Geographical Review* **38**(1): 55–94.

Tsakiris G, Vangelis H. 2004. Towards a drought watch system based on spatial SPI. *Water Resources Management* **18**(1): 1–12.

Ward JH. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**(301): 236–244.

Wells N, Goddard S, Hayes MJ. 2004. A self-calibrating Palmer drought severity index.*Journal of Climate* **15**: 2335–2351.

Wheaton EE, Arthur LM, Chorney B, Shewchuk S, Thorpe J, Whiting J, Wittrock V. 1992. The prairies drought of 1988. *Climatological Bulletin* **26**: 188–205.

Wheaton EE, Kulshreshtha S, Wittrock V, Koshida G. 2008. Dry times: hard lessons from the Canadian drought of 2001 and 2002. *The Canadian Geographer* **52**(2): 241–262.

Wheaton EE, Wittrock V, Kulshreshtha S, Koshida G, Grant C, Chipanshi A, Bonsal B. 2005. Lessons learned from the Canadian drought years of 2001 and 2002: synthesis report. Saskatchewan Research Council, Publication No. 11602-46E03.

Whitmore JS. 2000. Hardy and drought – evasive cereal crops. In *Drought Management on Farmland*, Whitmore JS (ed). Springer: London, 79–97.

Williams CL, Liebman M, Edwards JW, James DE, Singer JW, Arritt R, Herzmann D. 2008. Patterns of regional yield stability in association with regional environmental characteristics. *Crop Science* **48**(4): 1545–1559.

Wu H, Hubbard KG, Wilhite DA. 2004. An agricultural drought risk-assessment model for corn and soybeans. *International Journal of Climatology* **24**(6): 723–741.

