INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **19**: 1517–1534 (1999)

NORTH AMERICAN SNOW EXTENT: 1900–1994

ALLAN FREI[a,*], DAVID A. ROBINSON[b,1] and MARILYN G. HUGHES[b,2]
[a] *CIRES/NSIDC, University of Colorado, Campus Box 449, Boulder, CO 80309, USA*
[b] *Rutgers University, Department of Geography, New Brunswick, NJ 08903, USA*

Received 7 August 1997
Revised 11 March 1999
Accepted 13 March 1999

ABSTRACT

Historical fluctuations of North American snow extent from November through March are reconstructed back to 1900 using a combination of satellite and station observations. Using results of principal components analyses (PCA) from a companion study (Frei, A. and Robinson, D.A. *Int. J. Climatol.*, this volume), simple and multiple linear regression models are used to take advantage of the spatial coverage of satellite observations and the temporal extent of station observations. This analysis more than triples the remotely-sensed record length, which begins in 1972.

Model results indicate that North American winter snow extent tended to increase between the 1930s and around 1980, followed by a subsequent decrease during the 1980s. Long-term trends during November are less dramatic, with small increases since the 1960s. During March a different signal is observed, with snow extent decreasing since the 1950s. These results suggest a possible shift in the snow season.

Historical signals from smaller regions within North America are identified during December and January. During December, the continental-scale signal is driven mainly by fluctuations over the western US, while January fluctuations are more strongly driven by an eastern signal.

Models are sufficiently accurate to estimate changes in interannual variability over North America only during February, as well as over the eastern portion of the continent during December and January. Continental-scale interannual variability during February has been high since the mid-1970s compared to any previous time this decade. Regional-scale interannual variability over eastern North America in January has also been higher in recent years, but in December the highest interannual variability occurred during the 1940s. Copyright © 1999 Royal Meteorological Society.

KEY WORDS: snow; North America; climate variability; linear regression analysis; principal components analysis

1. INTRODUCTION

One of the most important challenges facing climatologists is the development of methods to assist in the detection of global climate fluctuations. The 'fingerprint' approach to climate change detection, as defined by a panel of experts convened by the US Department of Energy (Pennell *et al.*, 1993), involves a comparison of observations of one or more climate variables to a predicted temporal or regional pattern of effects from a known perturbation to the earth's climate system. Using this approach, a change in climate must be detectable, and must be attributable to a known cause, such as an increase in the concentration of atmospheric greenhouse gases. A priority for the implementation of this methodology is to estimate the natural variability of climate. Towards meeting this goal, in this analysis the historical variability of North American snow extent between November and March is reconstructed back to 1900, more than tripling the current satellite record.

A valid climate change detection scheme requires knowledge of historical as well as spatial variability (Goodison and Walker, 1993). Station observations have been used in several studies to examine

---

* Correspondence to: CIRES/NSIDC, University of Colorado, Campus Box 449, Boulder, CO 80309, USA. Tel.: +1 303 4926085; fax: +1 303 492 2468; e-mail: frei@kryos.colorado.edu
[1] E-mail: drobins@rci.rutgers.edu
[2] E-mail: mghughes@gandalf-rutgers.edu

CCC 0899–8418/99/141517–18$17.50

continental-scale historical variations in snow depth. Over Russia, the trend between 1936 and 1983 has been for increasing snow depths north of 60°N (Ye *et al.*, 1998) despite increasing temperatures in those regions (Fallot *et al.*, 1997). South of 60°N the trend has been for decreasing snow depths (Ye *et al.*, 1998). Over Canada, winter and early-spring snow depths decreased between 1946 and 1995, coinciding with temperature increases (Brown and Braaten, 1998).

Snow depth observations from meteorological stations can also be used to estimate historical fluctuations of snow extent. For example, Hughes and Robinson (1996) examined fluctuations of snow extent over the US Great Plains back to 1910. Brown and Goodison (1996) reconstructed historical snow cover back to 1915 for selected regions in southern Canada using a mass balance approach. Brown ('Northern Hemisphere snow cover variability and change, 1915–1997', *J. Clim.*, submitted) uses station observations of depth in a snow-index approach to estimate historical continental-scale fluctuations.

Others have used empirical orthogonal functions derived from satellite observations, in combination with pre-satellite era *in situ* observations, to estimate historical variations in other climate variables, such as sea-surface temperature (Smith *et al.*, 1996). Brown (1996, 1997) and Hughes *et al.* (1996) show that it is possible to reconstruct continental- to hemispheric-scale snow variability using historical station observations in conjunction with PCA. Such a technique applied in this analysis.

In this study, station observations of snow depth from across the continental US, plus a few stations from southern Canada, are combined with remotely-sensed observations of snow extent in such a way as to take advantage of the best qualities of each data set. Station records, going back almost 100 years in some cases, have the temporal component required for historical reconstruction, but are spatially incomplete. Each station observation of snow depth is valid only at the station site. In addition, the spatial distribution of stations is temporally inconsistent, and is not sufficiently dense as to cover the variety of terrain, vegetation, and land-use found across North America.

Satellite observations, on the other hand, are spatially integrated so that coverage across snow covered Northern Hemisphere lands is complete. However, reliable estimates of snow extent from satellites are available only since 1972. The satellite record indicates that during the period from the mid-1980s through the late 1990s, Northern Hemisphere annually averaged snow was significantly less extensive compared to the period from 1972 through the mid-1980s (Robinson, 1999). Such changes in snow extent could be interpreted as a decreasing trend, which might be considered to be consistent with theories of global warming. However, in the context of the ongoing debate regarding global climate change, it is important to obtain time series that are temporally as extensive and complete as possible.

In a companion study to this one (Frei and Robinson, 1999) a principal components analysis (PCA) of remotely-sensed observations of snow extent between 1972 and 1993 is used to identify key or 'coherent' (defined below) regions of snow extent fluctuations across the Northern Hemisphere. In this study, snow depth observations from 1900–1993 are used in conjunction with the results of Frei and Robinson (1999) to reconstruct regional- and continental-scale North American snow extent during the 20th century prior to the satellite era. Three properties of the observational data sets make this reconstruction possible.

First, regions of 'coherent' snow extent fluctuations over North America have been identified and described using PCA (Frei and Robinson, 1999). By 'coherent' it is meant that interannual fluctuations of snow extent over all grid cells within each region have high temporal correlations. Coherent regions are defined as having grid cells with loading factors ≥ 0.71 on to a common component and, therefore, have $\geq 50\%$ of their variance explained by a common score time series. Only those coherent regions located over the same geographic area as station observations (i.e. the continental US) can be included in this analysis. As a result, we are able to reconstruct historical snow extent between November and March; during other months, all or most coherent regions lie outside of the spatial domain of this analysis. Due to the coherency of the signal within each region, station observations can be used in simple linear regression models to estimate regional snow extent, even when relatively few stations are available within a region. This is especially important for estimating snow extent during the early part of this century, when fewer station observations are available.

The second property that makes this analysis possible is that much of the variance in continental-scale snow extent fluctuations is explained by signals from 'coherent' regions (for most months $\geq 60\%$) (Frei and Robinson, 1999). This property makes it possible to estimate North American snow extent using observations only from within coherent regions, which represent a small portion of the continental surface area.

Finally, the locations of coherent regions have been stable during the entire period of study. Frei and Robinson (1999) based their PCA on satellite observations from 1972 to 1994. It is uncertain *a priori* whether regions of coherent snow extent fluctuations had similar geographical boundaries during the pre-satellite era. In this paper we present results of a correlation analysis of historical station observations that shows that the geographic boundaries of coherent regions remained stable over time.

The technique used in this analysis has potential advantages as well as disadvantages. On the disadvantageous side, this method can potentially lead to underestimation of the magnitude of historical fluctuations for two reasons. First, regression models tend to underestimate variability. Second, if historical fluctuations in snow extent were of a sufficiently large magnitude so that the snow line during earlier times was either completely south or completely north of the key regions, then the historical reconstruction will capture the direction, but not the full magnitude, of the historical change. There are also advantages. First, for each time period, the maximum possible number of stations are included. Second, no filling or estimating of station values is required. Third, this method may be more applicable than alternatives over the Eurasian land mass, where a less dense station network is likely to be available, even when more station observations become available over currently data-sparse regions. Fourth, results of this analysis can be easily compared to future observations since satellite coverage will likely remain more continuous, compatible, and geographically complete than station-based data sets.

In the second section the observational data sets used in this analysis are described. The third section contains a brief presentation of the correlation analysis showing that the locations of coherent regions have remained stable since 1900. In the fourth section the methodology for constructing regional and continental-scale models is presented. The fifth section contains model results, and the sixth section contains a summary and conclusions.

2. OBSERVATIONAL DATA

2.1. Satellite observations

Charts of Northern Hemisphere snow cover extent, produced by the National Oceanic and Atmospheric Administration (NOAA), from January 1972 through December 1994 are used in this analysis. These weekly charts are derived from photographic copies of satellite images that are interpreted by trained meteorologists. The Very High Resolution Radiometer (VHRR) (1972–1978) and the Advanced VHRR (AVHRR) (1978–present), which are sensitive in visible wavelengths and provide spatial resolution of about 1 km, have been the main instruments used for snow cover detection. Although charts for the pre-VHRR period (1966–1972) period are available, snow extent was underestimated during those years, especially during autumn, due to problems with cloudiness (Wiesnet *et al*., 1987). Other limitations of visible imagery for snow cover detection include problems with low solar illumination, dense forest cover, and steep terrain. However, at a monthly resolution these data are suitable for climatic studies (Kukla and Robinson, 1981; Wiesnet *et al*., 1987).

The weekly charts are digitized using the National Meteorological Center Limited Area Fine Mesh Grid, an 89×89 cell cartesian grid overlying a Northern Hemisphere polar projection. This grid includes more than 5000 cells over land, with cell resolution between approximately 16000 km^2 and 42000 km^2. For each week, if a cell is interpreted to be $\geq 50\%$ snow covered, it is considered to be completely covered; otherwise it is charted as snow free. Weekly charts identify snow in a grid cell on the latest day of the week during which the ground was visible. Robinson (1993a) devised an improved routine to calculate monthly average snow frequencies from weekly charts, and applied corrections for inconsistencies in the demarcation of land versus ocean grid cells used in chart digitization. The data set used in this study contains monthly average snow cover frequencies for each grid cell for the 23-year period between 1972 and 1994.

2.2. *Station observations*

Station observations of snow cover over the contiguous 48 US states are obtained from the Historical Daily Climate Dataset (HDCD), which contains long-term digitized records of daily snow depth, snowfall, precipitation, and maximum and minimum temperatures from over 1100 cooperative climate stations (Robinson, 1993b). Snow cover records pre-date 1920 for approximately 45% of the stations, while all stations report snow measurements between 1948 and 1993. All data are highly quality controlled and checked for inconsistencies, errors, and missing values. Snow depth is reported in whole inches, with depths of less than 0.5 in. reported as a trace, and not recorded in the digitized HDCD files. In this study, station data from approximately 550 HDCD stations, each of which fall within at least one coherent region, and some of which have observations as far back as 1900, are utilized to estimate regional and continental snow extent. This data set is augmented by daily snow depth observations from 14 Canadian stations, for which digital records extend back to 1955.

3. STABILITY OF HISTORICAL REGIONS

A critical assumption in this analysis is that the coherent regions identified during the satellite era by Frei and Robinson (1999) remained stable during the earlier part of the century. Results of a correlation analysis of historical station observations, that was performed independently of the historical reconstruction, indicates this to be the case. Figure 1 shows locations of these regions for November through March. More details of this analysis can be found in Frei (1997).

A subset of the HDCD consisting of 204 stations that are well distributed across the continental US, and that have records extending prior to 1920, are selected for analysis. Monthly time series of snow cover duration anomalies (number of days per month with at least 1 in. of snow subtracted from the 1972–1993 mean) are computed for each station. Time series from these 204 stations are used to investigate coherent regions during three time periods: early (1900–1939), middle (1940–1971), and recent (1972–1993). For each coherent region, a regional-average snow duration time series is calculated during each time period. Then, Pearson correlation coefficients between the regionally-averaged time series and each individual station time series across the US are calculated. The correlation coefficients are contoured, and the isolines compared to the boundaries of coherent regions.

Boundaries associated with isolines are similar to those of coherent regions, indicating that the locations of coherent regions have been stable during the entire study period. For example, the correlation maps for December PC1 during the early (1900–1939), middle (1940–1971), and recent (1972–1993) time periods all show that western stations are significantly and positively correlated to each other and uncorrelated to eastern stations (Figure 2). Also shown in Figure 2 are historical correlations for March PC2. All regions except for one give positive results. The only example where the boundary of a coherent region during an earlier time period appears to be different than during the satellite era is January PC4 between 1900 and 1939. In this region, station observations are of poorest quality and it is, therefore, excluded from this historical analysis.

4. METHODOLOGY FOR MODELLING REGIONAL- AND CONTINENTAL-SCALE SNOW EXTENT

4.1. *Regional-scale model*

4.1.1. Model formulation. Simple linear regression models (Equation (1)) are used to estimate regional snow extent during the satellite and pre-satellite eras. The dependent variable is the PCA score time series of regional snow extent derived in Frei and Robinson (1999). The independent variable is regionally-averaged snow duration derived from station observations within the PC region. For each historical period (1900–1909, 1910–1919, ..., 1950–1959, 1960–1971, 1972–1993) a separate regression is

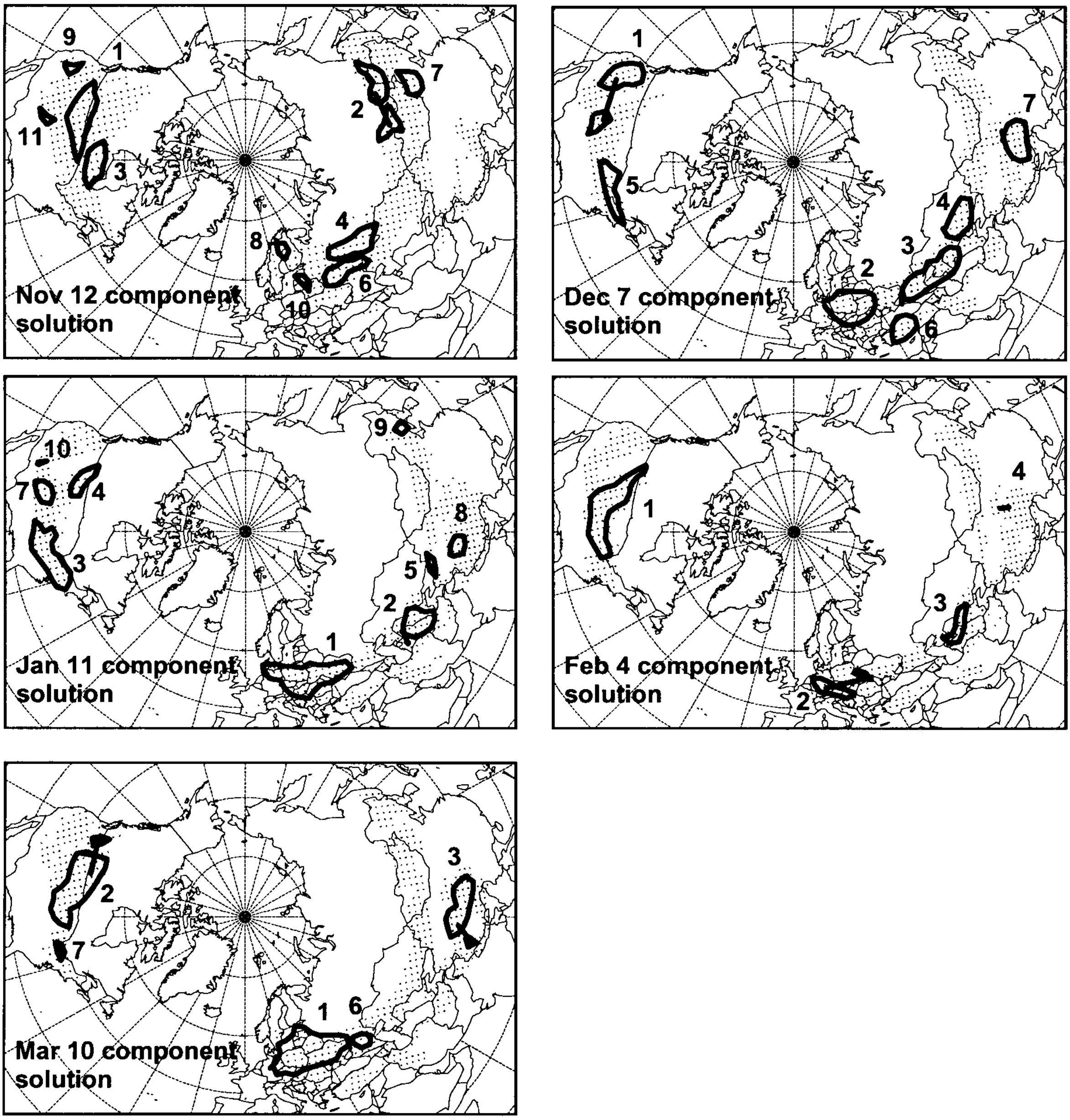

Figure 1. Key regions of snow extent fluctuations for November through March. Regions are derived from principal components analysis (PCA) of gridded Northern Hemisphere snow extent. Lines indicate 0.7 loading contour of grid cells, and numbers indicate principal component number in ascending order (PC1 explains most variance). Dots indicate 'active' grid cells, i.e. those that are included in the PCA

performed, so that our model includes the effect of the changing distribution of stations. Each regression model uses the station *distribution* from the historical time period, but is computed using *observations* from the satellite era (1972–1993) (this is described below in the section *Choosing stations for inclusion in regional average*). In the remainder of this paper we sometimes refer to the historical periods as 'decades', even though the 1960–1971 and 1972–1993 periods are not 10 years in length.

$$Y_i(t, x) = a_{ij} + b_{ij}x_{ij}(t) \pm \text{S.E.}_{ij} \qquad (1)$$

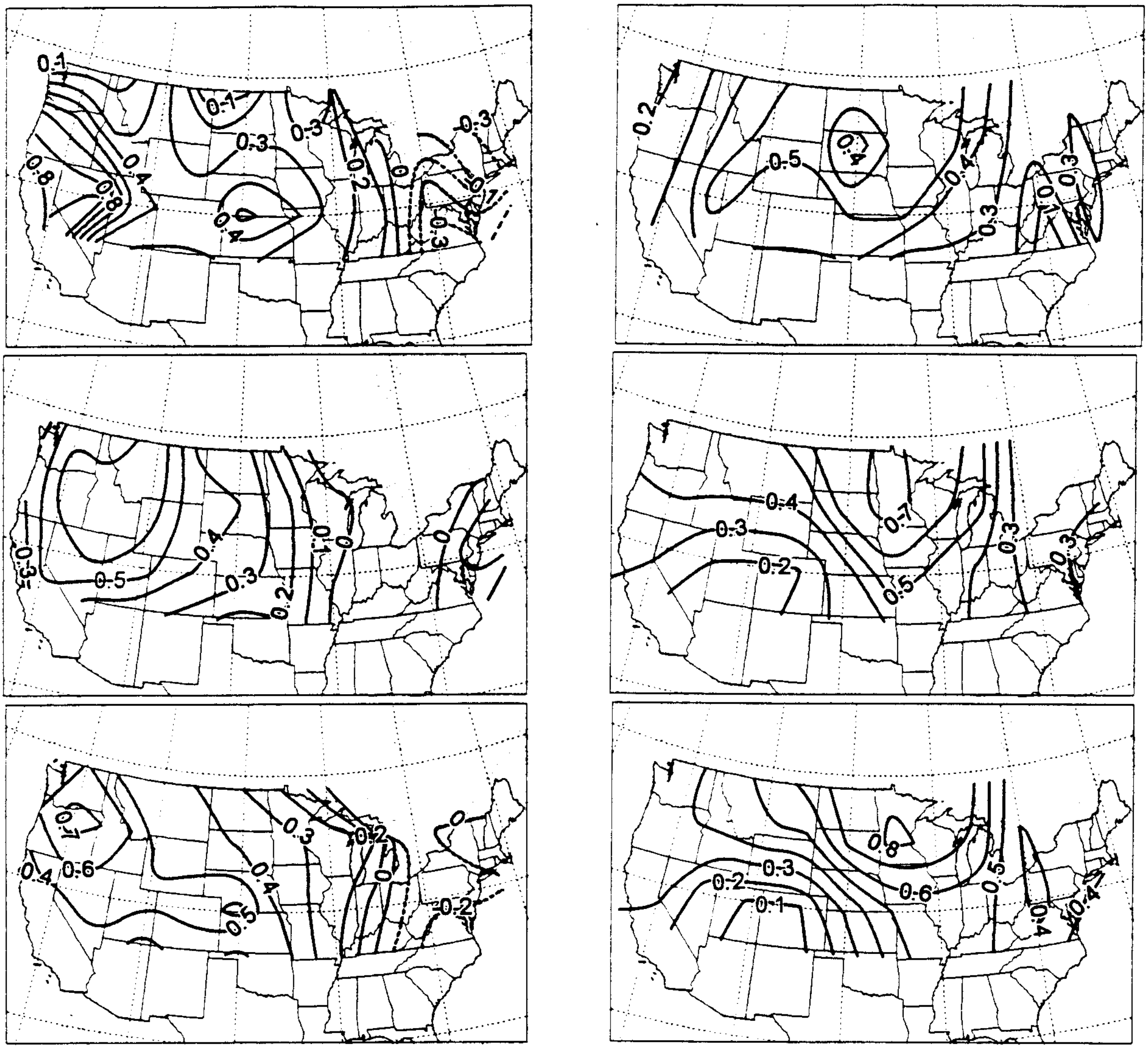

Figure 2. Contours of Pearson correlation coefficients between individual station time series and the average of all stations falling within the 0.7 loading contour of December PC1 (left) and March PC2 (right) for 1900–1939 (top), 1940–1971 (middle), and 1972–1993 (bottom)

where $Y = \text{PC}$ score time series of snow extent fluctuations for specified region (i) for time period 1972–1993; $x =$ station-based regional mean snow cover duration for 1972–1993 (# days per month with snow depth ≥ 2.54 cm) derived from station observations using station distribution from specified time period (j); $t =$ year (1972, 1973, ..., 1993); $a =$ regression constant; $b =$ regression coefficient; S.E. = standard error of prediction (calculated from cross-validated prediction); $i =$ PC region; and $j =$ time period for station distributions $j = 1$ (1900–1909), 2 (1910–1919), ..., 8 (1972–1993).

4.1.2. Calculation of station-based snow duration. For each station, a time series of monthly snow duration (# days per month with snow cover) for each month between November and March is computed from snow-depth observations. A minimum depth of 1 in. (2.54 cm) is required for a station to be considered snow covered, since that is approximately the minimum depth that NOAA satellite observations can identify (Kukla and Robinson, 1981). Months with ≥ 3 days of missing daily observations are considered missing.

4.1.3. Choosing stations for inclusion in regional average. Most stations have months with at least some missing values. For each time period, there are two criteria for inclusion of stations in the regression model. First, stations must have non-missing values for at least two-thirds of the years during the reconstruction period (7 years of 10, 8 years of 11, or 15 years of 22). Second, stations must have non-missing values during two-thirds of the years during the 1972–1993 period (15 of 22 years). This accounts for the effect of historical changes in station distributions.

For example, 39 stations are located within the boundaries of December PC1 that had non-missing values during 8 of 11 years from 1960–1971 and had non-missing values during 15 of 22 years from 1972–1993 (Table I). These stations constitute the station distribution for December PC1 for the 1960–1971 period. Only these stations are included in the regionally-averaged snow duration time series that is used in a regression against satellite observations from the period 1972–1993. The procedure is repeated using station distributions for each region during each time period. For most regions, the number of stations available is fewer than 20 during the 1910s, but then rises steadily until the satellite era when there are up to 110 stations per region. However, in general, only a few stations are required for regression results to be significant (see Results section).

4.1.4. Evaluation of regional models. Since long time series are not available for each model, it is impossible to train the model during one period and test it during another period. Therefore, in this analysis, evaluation of the model is performed using the cross-validation technique as described by Livezey (1995). In this technique, for each regression the coefficients are calculated n times ($n = \#$ years in regression), each time removing one point (i.e. year) from the model. Thus, the predicted value for each time point is calculated from a regression equation using the $(n-1)$ *other* time points. Then, the cross-validated percent of variance explained ($xr\,^\wedge 2$) is calculated using cross-validated predicted values, as in Michaelson (1987) (Equation (2)). S.E. $(x-se)$ is similarly calculated (Equation (3)).

$$xr\,^\wedge 2 = 1 - xsse/ssa. \tag{2}$$

$$xse = \sqrt{xsse/(n-2)}. \tag{3}$$

In Equations (2) and (3), *xsse* is the cross-validated sum of square errors,

Table I. Number of stations available for regression models

Region	1972–1993	1960–1971	1950–1959	1940–1949	1930–1939	1920–1929	1910–1919	1900–1909
Nov1	**48**	**46**	**32**	**14**	**14**	**10**	**12**	2
Nov9	7	6	6	5	5	1	0	0
Nov11	11	11	9	7	4	4	3	1
Dec1	**45**	**39**	**32**	**32**	**29**	**8**	**7**	**5**
Dec5	**31**	**29**	**26**	**20**	**13**	**7**	**6**	**5**
Jan3	**89**	**85**	**75**	**65**	**57**	**22**	**16**	**13**
Jan4	13	9	6	2	1	1	2	0
Jan7	**18**	**17**	**13**	**10**	**8**	**8**	**5**	**4**
Feb1	**111**	**99**	**80**	**66**	**48**	**45**	**37**	**28**
Mar2	**74**	**68**	**51**	**27**	**23**	**19**	**17**	12
Mar7	5	5	5	3	3	2	2	1
Apr3	11	11	4	0	1	1	0	0

For each region and each time period, this table indicates the number of stations included in historical reconstruction of regional snow extent. See text for explanation of criteria for inclusion. Bold indicates which regression models were used in reconstructions of regional snow extent.

$$xsse = \sum (xpredicted - y)^2$$

and *ssa* is the sum of square anomalies,

$$ssa = \sum (y - \bar{y})^2.$$

Results of regression Equation (1) indicate that station-based snow duration anomalies are accurate indicators of regional snow extent (Table II). Cross-validated correlation coefficients ($x - r \wedge 2$) are calculated for each region/decade. Regressions are significant for almost all regions during most time periods. Values shown in bold in Table II are those with $x - r \wedge 2 \geq 0.50$, indicating that station observations are sufficient to explain at least 50% of the variance in the remotely-sensed snow extent record. These coincide with the regions for which a greater number of stations are available (Table I). Also shown in Table II are the differences between the standard $r \wedge 2$ and cross-validated $xr \wedge 2$ values. The difference gives an indication of the magnitude of overestimation of explained variance given by the standard $r \wedge 2$ value. For those regions/time periods with generally higher $r \wedge 2$ values, cross-validation affects the estimate of explained variance by < 0.1 (sometimes much less); while for those with lower values, the difference tends to be > 0.1. This indicates that the models are robust for regions with higher $r \wedge 2$ values. Bold values in Table II have $xr \wedge 2 > 0.5$, and indicate those models for which historical reconstructions are performed.

The regression models are also robust in the sense that whenever the model estimates a significantly positive or negative value, the observed value is of the same sign. In this case the criteria for significance is whether the predicted value is at least 1 (cross-validated) S.E. from the mean.

4.2. Continental-scale model

4.2.1. Model formulation. Stepwise multiple linear regression analysis is used to estimate continental-scale snow extent. North American snow extent is the dependent variable, and station-based regional snow duration time series from each region are the predictors. As described earlier in this paper, the effects of changing station distributions are considered in this model. The stepwise procedure is performed so that if a regional time series does not contribute significantly ($p = 0.05$) to the regression, that regional time series is removed from the model. The model equation is:

$$NA(t, x) = a_j + \sum_{i=1}^{n} [b_{ij} x_{ij}(t)] \pm \text{S.E.}_{,j} \tag{4}$$

where NA = remotely-sensed North American snow extent (10^6 km^2); x = station-based regional snow cover duration anomaly (# days per month with snow depth 2.54 cm); t = year (1972, 1973, ..., 1993); b = regression coefficient; i = PC region; n = number of PC regions used in multiple regression; S.E. = standard error of prediction; and j = decadally specific station distributions prior to the satellite era $j = 1$ (1900–1909), 2 (1910–1919), ..., 8 (1972–1993).

4.2.2. Evaluation of North American model. As with regional models, cross-validated techniques are used to evaluate continental models. Cross validated $xr \wedge 2$ values (Table III) indicate that station observations can be used to estimate continental-scale snow extent fluctuations. The best results are found in February, and the worst in January. $xr \wedge 2$ values for continental snow extent during the satellite era (1972–1993) range from 0.465 (January) to 0.807 (February). The difference between cross-validated and standard $r \wedge 2$ values is greatest during January and smallest during February, indicating that reconstructions of snow extent for February are more stable than for other months. For months other than January, most difference values are < 0.1, and the maximum is 0.11. During January, the differences are between 0.14 and 0.15 since 1930, and as high as 0.18 during earlier decades. In order to account for the changing accuracy of the regression models, in the *model results* section cross-validated S.E. (*xse*) values for each decade are used as confidence limits.

Table II. Evaluation of simple linear regression models of regional snow extent

Region		1972–1993	1960–1971	1950–1959	1940–1949	1930–1939	1920–1929	1910–1919	1900–1909
Nov1	*xr* ^ 2	**0.744**	**0.743**	**0.770**	**0.740**	**0.768**	**0.703**	**0.699**	na
	r ^ 2	0.788	0.787	0.807	0.811	0.823	0.763	0.772	na
	diff	0.044	0.044	0.037	0.071	0.055	0.060	0.073	na
Nov9	*xr* ^ 2	0.133	0.025*	0.133	0.096*	0.096*	na	na	na
	r ^ 2	0.259	0.259	0.250	0.241	0.241	na	na	na
	diff	0.126	0.234	0.117	0.145	0.145	na	na	na
Nov11	*xr* ^ 2	0.208	0.208	0.284	0.384	−0.121*	−0.121*	−0.186*	0.097*
	r ^ 2	0.467	0.467	0.480	0.510	0.339	0.339	0.318	0.328
	diff	0.259	0.259	0.196	0.126	0.460	0.460	0.504	0.231
Dec1	*xr* ^ 2	**0.654**	**0.660**	**0.597**	**0.583**	**0.586**	**0.511**	**0.526**	**0.569**
	r ^ 2	0.711	0.720	0.671	0.664	0.662	0.592	0.598	0.636
	diff	0.057	0.060	0.074	0.081	0.076	0.081	0.072	0.067
Dec5	*xr* ^ 2	**0.855**	**0.845**	**0.819**	**0.812**	**0.715**	**0.599**	**0.548**	**0.660**
	r ^ 2	0.879	0.870	0.849	0.843	0.759	0.664	0.614	0.709
	diff	0.024	0.025	0.030	0.031	0.044	0.065	0.066	0.049
Jan3	*xr* ^ 2	**0.777**	**0.777**	**0.773**	**0.778**	**0.785**	**0.684**	**0.647**	**0.674**
	r ^ 2	0.816	0.813	0.812	0.816	0.824	0.743	0.714	0.729
	diff	0.039	0.036	0.039	0.038	0.039	0.059	0.067	0.055
Jan4	*xr* ^ 2	0.246	0.250	0.196	na	na	na	na	na
	r ^ 2	0.369	0.369	0.324	na	na	na	na	na
	diff	0.123	0.119	0.128	na	na	na	na	na
Jan7	*xr* ^ 2	**0.695**	**0.678**	**0.677**	**0.644**	**0.648**	**0.648**	**0.573**	**0.553**
	r ^ 2	0.754	0.737	0.744	0.718	0.721	0.721	0.641	0.626
	diff	0.059	0.059	0.067	0.074	0.073	0.073	0.068	0.073
Feb1	*xr* ^ 2	**0.832**	**0.833**	**0.818**	**0.809**	**0.818**	**0.804**	**0.816**	**0.805**
	r ^ 2	0.854	0.855	0.843	0.834	0.843	0.830	0.840	0.831
	diff	0.022	0.022	0.025	0.025	0.025	0.026	0.024	0.026
Mar2	*xr* ^ 2	**0.696**	**0.688**	**0.679**	**0.668**	**0.589**	**0.614**	**0.623**	**0.456**
	r ^ 2	0.748	0.743	0.735	0.727	0.667	0.687	0.693	0.564
	diff	0.052	0.055	0.056	0.059	0.078	0.073	0.070	0.108
Mar7	*xr* ^ 2	0.334	0.334	0.334	0.288	0.288	0.228	0.228	na
	r ^ 2	0.425	0.425	0.425	0.384	0.384	0.340	0.340	na
	diff	0.091	0.091	0.091	0.096	0.096	0.112	0.112	na
Apr3	*xr* ^ 2	0.225	0.225	0.027*	na	na	na	na	na
	r ^ 2	0.361	0.361	0.208	na	na	na	na	na
	diff	0.136	0.136	0.181	na	na	na	na	na

* Not significant (1-tailed; $p = 0.05$; $n = 22$).

xr ^ 2 = cross validated explained variance; *r* ^ 2 = linear regression explained variance; diff = difference between *r* ^ 2 and *xr* ^ 2, indicates the overestimation of explained variance by using standard, not cross-validated, values; na = not available for regression due to insufficient data. Bold indicates which regression models were used in reconstructions of regional snow extent.

5. MODEL RESULTS

In this section we present continental- and regional-scale reconstructions of snow extent. The models are sufficiently accurate to estimate decadal-scale fluctuations in North American snow extent during all months between November and March. Regional results are presented for December and January, months for which the stepwise multiple regression procedure identified two separate regions as contributing in a statistically significant way to the continental signal.

In addition, an evaluation of these models' abilities to estimate historical interannual variability is presented. These models are less suitable for estimating interannual variability than for estimating yearly values. Nevertheless, in limited cases they provide sufficient information to estimate historical variability.

5.1. Continental snow extent

Continental snow extent prior to the satellite era tended to be lower than during the satellite era. From November through February, many reconstructed values prior to 1972 for continental snow extent tend to be significantly below the 1972–1993 mean, while few values are significantly above. For November, five values between 1900 and 1971 are significantly above the mean, while 15 are below. During December, January, and February, only two, one, and two values, respectively, are significantly above, while 23, 32, and 41 fall below.

During March, between 1900 and 1971 ten values fall above, and 15 below, the 1972–1993 mean. However, they are not evenly distributed. We observe a period of increasing snow extent between the late 1920s and around 1950: all the low values in March occur prior to 1943, and all the higher values occur between 1948 and 1969. Thus, the historical record of snow extent is different for March than for other months.

Figure 3 shows the reconstructed time series of historical North American snow extent for November through March. In this discussion of long-term trends, we focus on the 9-year running mean values to characterize decadal-scale fluctuations. During November, the month with the smallest decadal-scale variability, snow extent fluctuated between 10.75 and 11.25×10^6 km^2 for most of the century. There is some indication of increasing values during the last three decades of the record, with the highest smoothed values occurring at the end of the record during the 1980s. This trend continues past 1994 (not shown in Figure 3): observed values during the 1990–1996 period were on average higher than any earlier comparable period. However, 1997 and 1998 snow extent during November were average and below average, respectively.

Table III. Evaluation of stepwise multiple regression models of North American snow extent

Month		1972–1993	1960–1971	1950–1959	1940–1949	1930–1939	1920–1929	1910–1919	1900–1909
November	*xr* ^ 2	0.488	0.488	0.481	0.429	0.440	0.476	0.441	na
	r ^ 2	0.575	0.575	0.565	0.518	0.528	0.551	0.542	na
	diff	0.087	0.087	0.084	0.089	0.088	0.075	0.101	na
December	*xr* ^ 2	0.634	0.644	0.535	0.551	0.554	0.562*	0.533*	0.567*
	r ^ 2	0.728	0.737	0.673	0.662	0.661	0.630*	0.608*	0.641*
	diff	0.094	0.093	0.138	0.111	0.107	0.068	0.075	0.074
January	*xr* ^ 2	0.465	0.468	0.437	0.461	0.437	0.422	0.362	0.386
	r ^ 2	0.607	0.609	0.587	0.608	0.589	0.580	0.543	0.549
	diff	0.142	0.141	0.150	0.147	0.152	0.158	0.181	0.163
February	*xr* ^ 2	0.807	0.813	0.824	0.832	0.841	0.832	0.824	0.820
	r ^ 2	0.832	0.838	0.847	0.853	0.862	0.854	0.846	0.844
	diff	0.025	0.025	0.023	0.021	0.021	0.022	0.022	0.024
March	*xr* ^ 2	0.556	0.541	0.553	0.610	0.534	0.586	0.614	0.521
	r ^ 2	0.622	0.609	0.620	0.669	0.610	0.652	0.669	0.605
	diff	0.066	0.068	0.067	0.059	0.076	0.066	0.055	0.084

All values are significant (1-tailed; $p = 0.05$; $n = 22$). * During the period 1900–1929, December PC5 became insignificant ($p = 0.5$), and was, therefore, removed from the model.

Regions included in historical reconstructions are: November, PCI only; December PCI and *PC5; January PC3 and PC7; February PC1 only; March PC2 only. *xr* ^ 2 = cross-validated explained variance; *r* ^ 2 = linear regression explained variance; diff = difference between *r* ^ 2 and *xr* ^ 2, indicates the overestimation of explained variance by using standard, not cross-validated, values; na = not available for regression due to insufficient data.

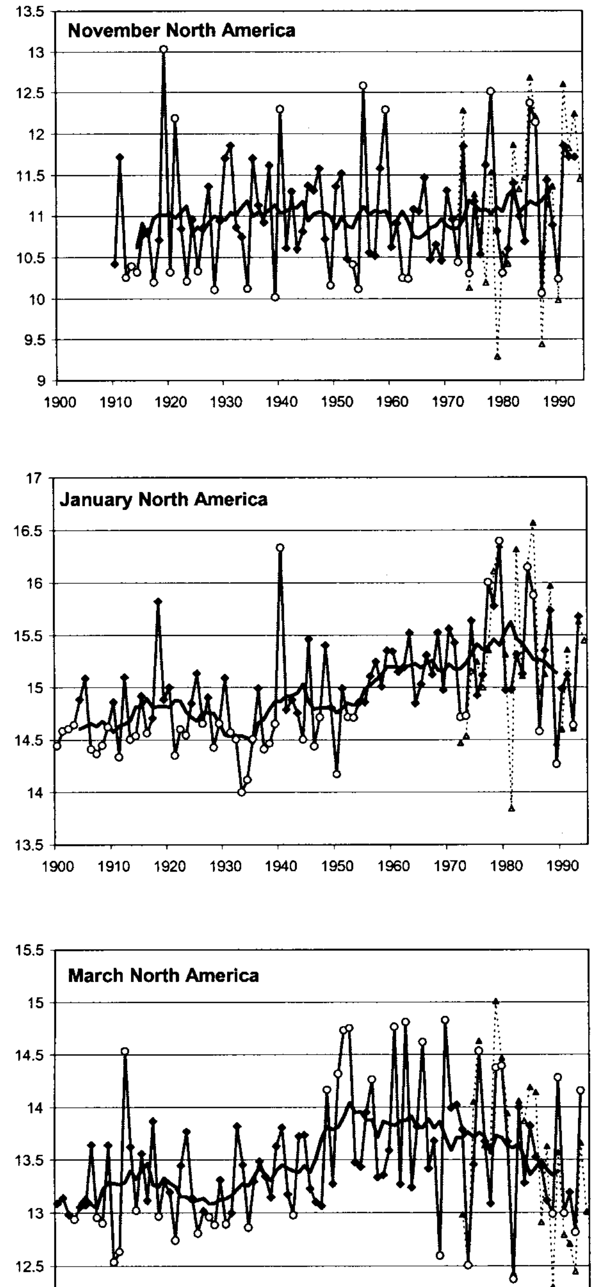

Figure 3. Reconstructed (1900–1993) and observed (1972–1994) North American snow extent from November through March. Units are 10^6 km^2. Solid line/diamonds are reconstructed values. Reconstructed points shown with circles are at least ± 1 standard error from 1972–1993 mean. Heavy black line is 9-year running mean of reconstructed time series. Dashed line is the observed time series

During December, running-mean values decreased from around 14×10^6 km^2 to 13.5×10^6 km^2 between the 1910s and the 1930s. Since that time, we see increasing snow extent, up to around 14×10^6 km^2 during the 1950s, 14.4×10^6 km^2 during the late 1960s, and 14.5×10^6 km^2 during the peak years of the 1980s. Brief periods of decreasing values are observed during the mid-1950s, the mid-1970s, and late

1980s. During the mid-1990s (not shown), December snow extent has remained above 14 × 10^6 km^2, except for 1998.

Snow extent during both January and February has tended to increase since around 1930, reaching peak values around 1980, and then falling during the 1980s. During the mid 1990s (not shown) snow extent has tended to decrease during January and February, reaching the second lowest values of the satellite era during 1998. In January, running-mean snow extent decreased from around 14.9 × 10^6 km^2 to 14.5 × 10^6 km^2 between the late 1910s and the early 1930s. After the early 1930s we observe an increase to around 15.25 × 10^6 km^2 during the 1960s and to over 15.5 × 10^6 km^2 in the early 1980s. Since the early 1980s, the running mean decreased by approximately 0.5 × 10^6 km^2. February snow extent dropped from around 14.5 × 10^6 km^2 during the early 1910s to 14 × 10^6 km^2 during the 1920s, and subsequently increased back to 14.5 × 10^6 km^2 during the 1940s and up to around 15 × 10^6 km^2 during around 1960s. After a drop of ~ 0.3 × 10^6 km^2 between 1960 and 1970, February snow extent increased dramatically to its peak values of the century of almost 15.5 × 10^6 km^2 during the early 1980s. As during January, February snow extent decreased between the early and late 1980s, in this case by ~ 0.7 × 10^6 km^2. March snow extent peaked around 1950 at 14 × 10^6 km^2, and dropped to below 13.5 × 10^6 km^2 during the 1980s. During the mid 1990s (not shown) March snow extent was comparable to the 1980s.

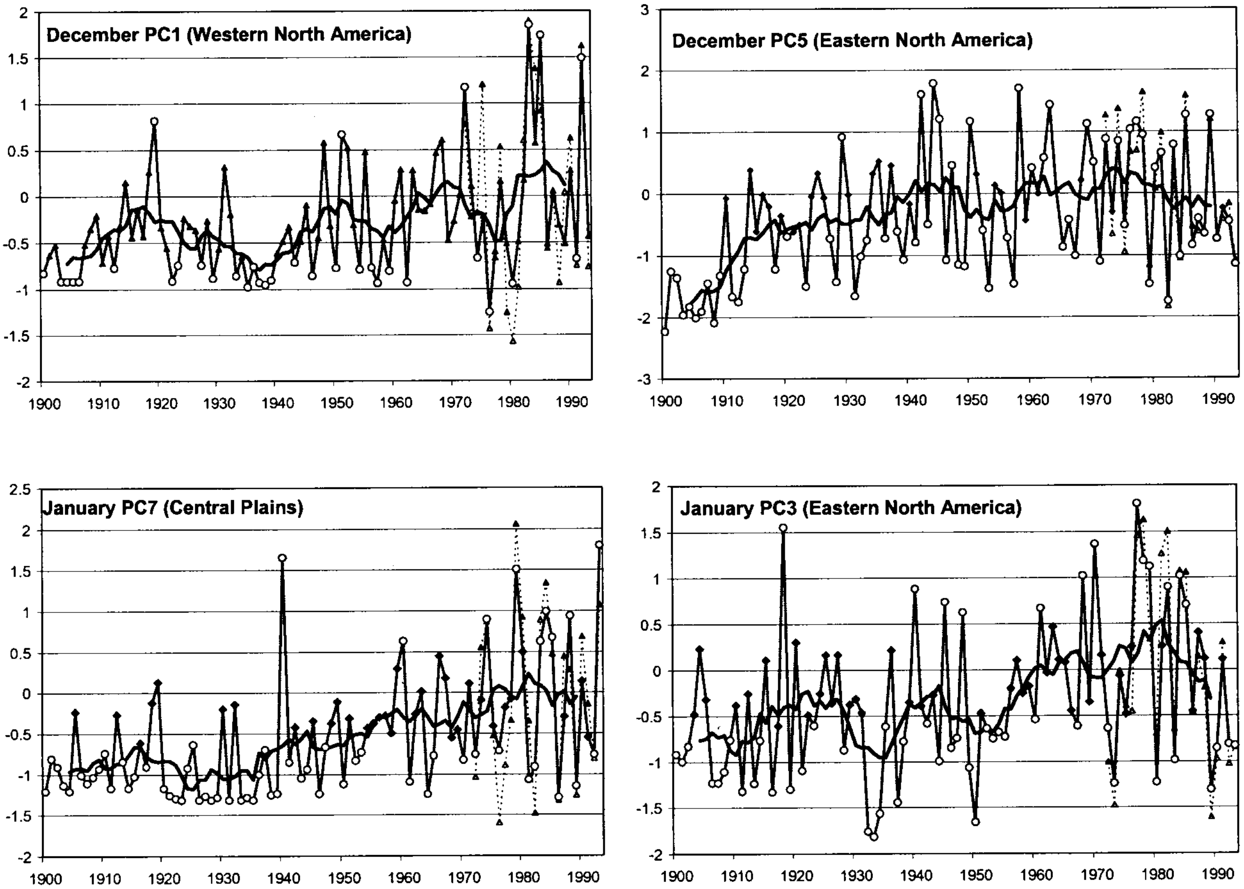

Figure 4. Reconstructed (1900–1993) and observed (1972–1994) regional snow extent during December and January. Units are normalized (standard deviations) derived from score time series of principal components analysis (PCA). Solid line/diamonds are reconstructed values. Reconstructed points shown with circles are at least ± 1 standard error from 1972–1993 mean. Heavy black line is 9-year running mean of reconstructed time series. Dashed line is the PCA score time series derived from observations

5.1.1. Regional snow extent. Only during the months of December and January are reconstructions possible for smaller regions within North America (Figure 4). During December over the western region, which is primarily driving the continental signal, historical variations are similar to those described above for the continent as a whole, with generally increasing snow extent between the 1930s and the 1980s. Over the east, a different signal emerges, where snow extent increases between 1900 and 1940, most steeply during the first two decades of the century. Since that time the signal has fluctuated around the 1972–1993 mean.

During January, the region most strongly influencing the continental signal is over the east. Here, during the first 50 years of the century we see large decadal-scale fluctuations about a mean that is below the 1972–1993 mean. Thus, during those first five decades, only 4 years had values > 1 *xse* above the 1972–1993 mean, while 27 years were significantly below. Subsequent to 1950, the running mean tends to rise for the next three decades and reach its peak values around 1980. Over the second January region—the central Plains—snow extent was consistently and extremely low until around 1940. At that time the signal began to gradually rise until around 1980. Decadal-scale fluctuations in the Plains are of smaller magnitude than over the east.

5.1.2. Continental interannual variability. Temporal changes in interannual variability are evaluated by examining the 9-year running mean of Inter-Quartile Range (IQR). A visual comparison of the reconstructed versus observed time series for each month (Figure 5) indicates that the models do a poor job of estimating interannual variability. However, in February, with greater than 80% of the observed variance captured by the models, it appears that the reconstruction is sufficiently accurate to estimate historical variability. The regression coefficients remain very stable from decade to decade (Table IV), indicating that changes in variability are not statistically induced. The predicted time series shown for November in Figure 5 also appears to mimic observed variations, even though $xr\,^\wedge 2$ values during November are much lower than during February. We consider the apparent success during November to be fortuitous, and conclude that for estimating historical changes in interannual variability of North American snow extent, the only model that provides a sufficiently accurate reconstruction is the month of February.

During February, interannual variability during the first five decades of the century tended to be low, with the 9-year running mean IQR fluctuating between around 0.2×10^6 km^2 and 0.6×10^6 km^2. Variability increased in the 1950s to almost 1×10^6 km^2, then decreased to around 0.4×10^6 km^2 during the mid-1960s. From the mid-1960s to the mid-1970s interannual variability rose steeply to greater than 1.2×10^6 km^2. Variability apparently decreased somewhat in the 1980s, but never back to the pre-1950 levels. As a result, the period of highest interannual variability occurs during the more recent portion of the record—since the mid-1970s.

5.1.3. Regional interannual variability. Most regional estimates of interannual variability are as poor as continental estimates, but there are three exceptions (Figure 6). Reconstructions for December PC5 (eastern North America), January PC3 (eastern North America), and February PC1 (central North America) all have $xr\,^\wedge 2$ values of approximately 0.8 (although February has the highest values), and appear to mimic observed interannual variability with some success. Since February PC1 controls continental-scale fluctuations, which were discussed above, we will discuss that region no further. The models do provide information about historical interannual variability during December and January over the east, but we have less confidence in these results than for February. The regression coefficients for December PC5 and January PC3 (Table IV) decrease prior to 1930, indicating that estimated changes in variability prior to the 1930s should be viewed with caution.

During December over the eastern portion of the continent, it appears that interannual variability was highest during the 1940s with IQR values of $\geq 2 \times 10^6$ km^2. The 1980s seem to have been the second highest period, followed by both the 1970s and 1930s. In January over the east, the period

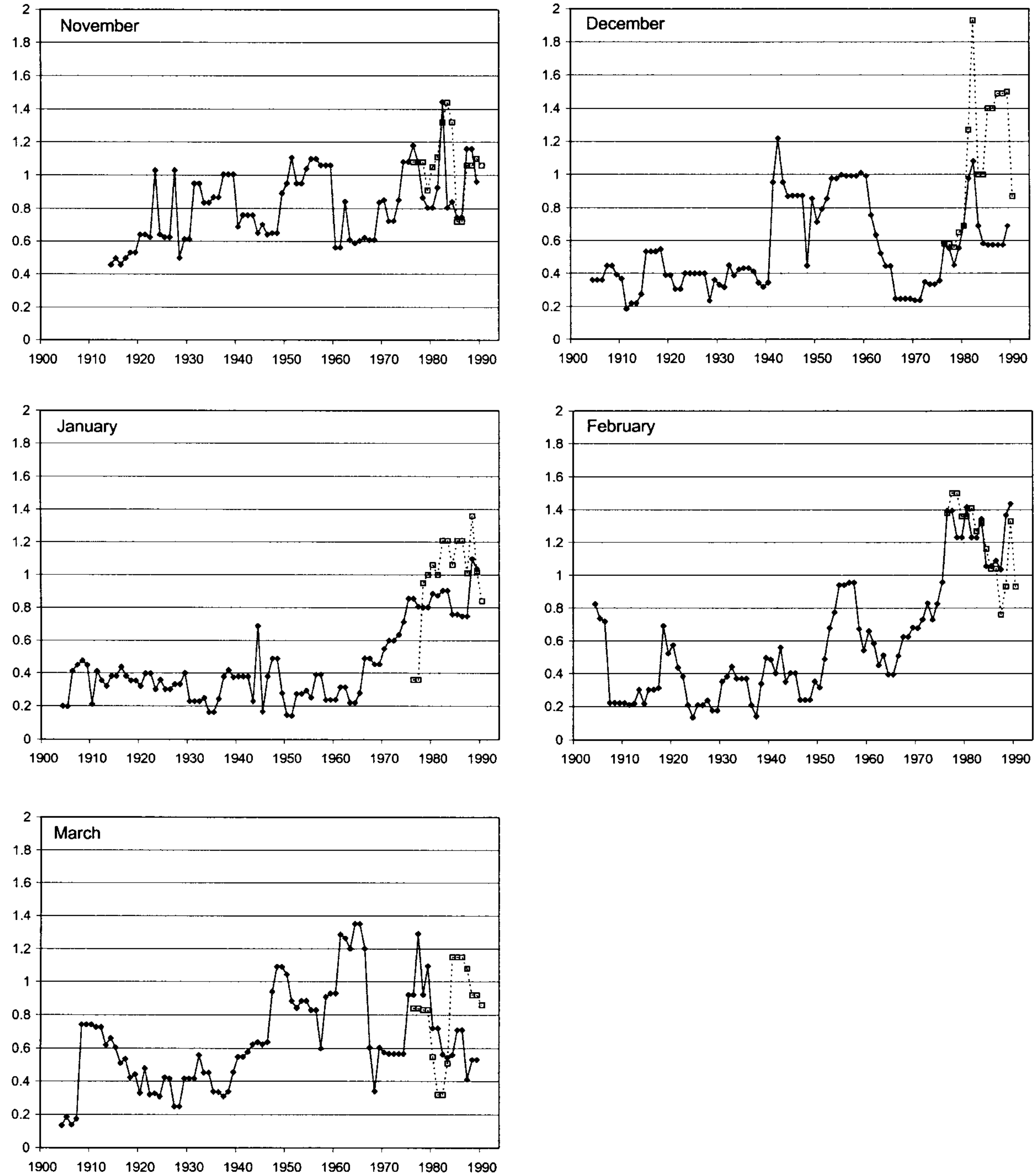

Figure 5. Nine-year running IQR in North American snow extent. Units are 10^6 km^2. Solid line/diamonds are predicted values from model, dashed lines/squares are observed

since the mid-1970s appears to have had higher interannual variability than earlier decades, although the 1930s (and late 1910s apparently) also had high variability. Particularly low interannual variability is observed during the late 1920s and early 1950s.

6. SUMMARY AND CONCLUSIONS

Station observations of snow depth, which have limited geographic coverage but are available back to the early 1900s, have been combined with satellite observations of snow extent, which provide comprehensive geographic coverage but are temporally limited to the period since 1972. Using principal components and regression analyses, fluctuations in snow extent across North America are reconstructed on regional- and continental-scales back to 1900, more than tripling the length of the satellite-era record, and enabling continental-scale fluctuations of snow extent to be viewed in the context of century-long fluctuations. As a result, we conclude the following.

6.1. (1) North American snow extent

Model results indicate a decrease in spring snow extent during the last three decades, and possibly a seasonal shift in the snow season. From December through March, snow tended to become more extensive between the 1930s and 1960s. Since the 1960s, November snow extent has tended to increase while March snow extent has decreased. These results are in agreement with those of Brown (1997, submitted), who also finds indications of decreasing spring snow extent over Eurasia.

6.2. (2) Regional snow extent

Differences in historical signals from regions within North America are identified during December and January. During December, the continental-scale signal is driven mainly by fluctuations over the western US, while we see a different signal over the east. January fluctuations are mostly driven by an eastern signal, and secondarily driven by a signal over the central Plains of the US.

6.3. (3) Interannual variability

Only during February are historical fluctuations in continental-scale snow extent reconstructed with sufficient accuracy to estimate changes in interannual variability; for regional snow extent, we have sufficient confidence in model results to estimate variability changes only over the east during December and January. During February, the period since the mid-1970s has had the highest interannual variability of the century. Similarly, in January over the east the recent period has had the highest interannual variability. However, in December, the eastern signal had higher interannual variability during the 1940s than during the recent era. We assign less confidence to the December and January regional results than to the February, continent-wide estimations of variability.

Table IV. Regression constants, coefficients, cross-validated S.E.s for selected models

Years	December PC5			January PC3			February PC1			February NA		
	a	*b*	*xse*	*a*	*b*	*xse*	*a*	*b*	*xse*	*a*	*b*	*xse*
1972–1993	−2.975	0.169	0.399	−1.980	0.132	0.485	−1.687	0.140	0.428	13.441	0.125	0.414
1960–1971	−2.862	0.167	0.413	−1.988	0.133	0.490	−1.672	0.140	0.427	13.451	0.125	0.408
1950–1959	−2.911	0.168	0.446	−2.006	0.133	0.489	−1.636	0.139	0.445	13.464	0.126	0.396
1940–1949	−2.707	0.162	0.455	−1.945	0.135	0.483	−1.613	0.139	0.457	13.472	0.127	0.387
1930–1939	−2.687	0.158	0.560	−1.962	0.137	0.476	−1.595	0.138	0.446	13.489	0.126	0.376
1920–1929	−2.032	0.129	0.664	−1.628	0.118	0.577	−1.610	0.137	0.462	13.470	0.126	0.386
1910–1919	−2.001	0.127	0.705	−1.617	0.118	0.610	−1.549	0.138	0.448	13.541	0.125	0.396
1900–1909	−2.468	0.135	0.611	−1.583	0.115	0.586	−1.554	0.136	0.461	13.530	0.124	0.400

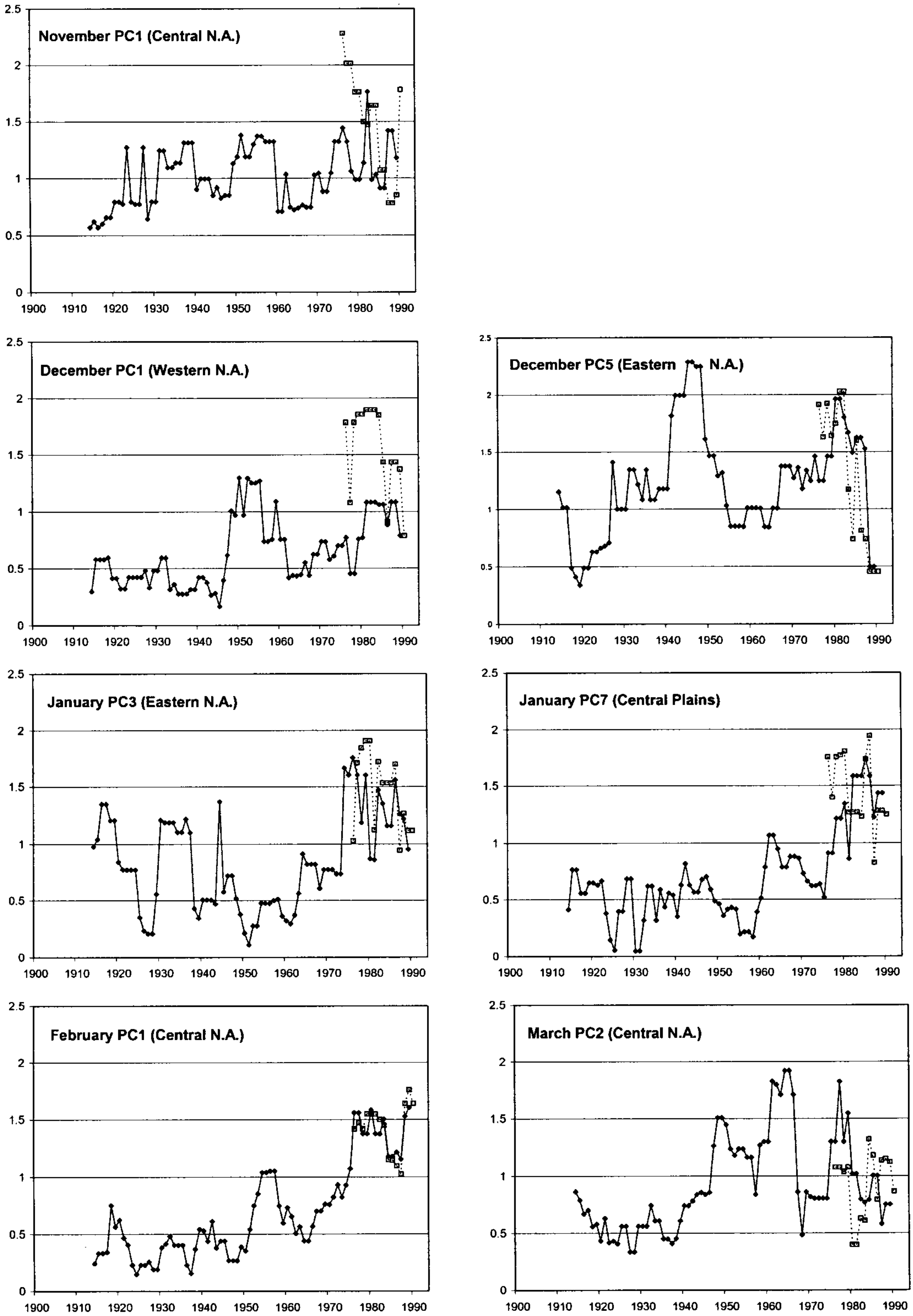

Figure 6. Nine-year running IQR in regional snow extent. Units are 10^6 km^2. Solid line/diamonds are predicted values from model, dashed lines/squares are observed

6.4. (4) North American snow extent and climate change

Current knowledge of historical fluctuations in snow extent is insufficient to identify the relationship between anthropogenic forcing and observed changes. The most intuitive expectation regarding climate change and snow extent is that the area covered by snow should tend to decrease as globally-averaged temperatures increase. While we have identified shifts in North American snow extent, only during spring has it decreased significantly. However, two vital pieces of information are missing. First, information about historical fluctuations over Eurasia is currently unavailable. Second, a simple linear decrease in continental- or global-scale snow extent may not, in fact, be the response of the climate system to increasing atmospheric greenhouse gases. The transient response of the cryosphere may have geographical and temporal patterns associated with changing atmospheric circulation patterns that are not intuitively obvious. Climate models must be evaluated and improved with regards to simulating snow extent (e.g. Frei and Robinson, 1998) and cryospheric processes in general. These two issues—the reconstruction of historical Eurasian snow extent, and the evaluation of snow simulations by climate models—are the focus of ongoing research efforts. Brown (1997, submitted) finds evidence of decreasing spring snow extent and depth over Eurasia as well as North America. The evidence to attribute observed changes in snow to either anthropogenic or natural factors is thus far inconclusive.

ACKNOWLEDGEMENTS

AF performed this work under a NASA Graduate Student Fellowship in Global Change Research and as a visiting fellow at CIRES/NSIDC, University of Colorado; DAR is supported by NSF grants ATM-9314721 and SBR-9320786, and NASA grant NAGW-3586. We thank Don Garrett at the NOAA Climate Analysis Center for providing satellite data, and Ross Brown, Robert Livezey, and one anonymous reviewer for helpful comments on this manuscript.

REFERENCES

Brown, R.D. 1996. 'Evaluation of methods for climatological reconstruction of snow depth and snow cover duration at Canadian meteorological stations', *Proceedings 53rd Eastern Snow Conference*, Williamsburg, VA, pp. 55–65.

Brown, R.D. 1997. 'Historical variability in Northern Hemisphere spring snow covered area', *Ann Glaciol.*, **25**, 340–346.

Brown, R.D. and Goodison, B.E. 1996. 'Interannual variability in reconstructed Canadian snow cover, 1915–1992', *J Climate*, **9**, 1299–1318.

Brown, R.D. and Braaten, R.O. 1998. 'Spatial and temporal variability of Canadian monthly snow depths, 1946–1995', *Atm-Ocean*, **36**, 37–45.

Fallot, J.M., Barry, R.G. and Hoogstrate, D. 1997. 'Variations in mean cold season temperature, precipitation and snow depths during the last 100 years in the former Soviet Union (FSU)', *Hydrol. Sci.—J. Sci. Hydrol.*, **42**(3), 301–327.

Frei, A. 1997. *Towards a Snow Cover Fingerprint for Climate Change Detection*, PhD Dissertation submitted to the Graduate School—New Brunswick, Rutgers, The State University of New Jersey, USA.

Frei, A. and Robinson, D.A. 1998. 'Evaluation of snow extent and its variability in the atmospheric model intercomparison project', *J. Geophys. Res.–Atmos.*, **103**(D8), 8859–8871.

Frei, A. and Robinson, D.A. 1999. 'Northern Hemisphere snow extent: regional variability 1972–1994', *Int. J. Climatol.*, this volume.

Goodison, B. and Walker, A.E. 1993. 'Data needs for detection of climate change: some challenges', *Snow Watch '92: Detection Strategies for Snow and Ice*, Glaciological Data Report GD-25, National Oceanic and Atmospheric Administration, pp. 241–253.

Hughes, M.G. and Robinson, D.A. 1996. 'Historical snow cover variability in the Great Plains region of the United States: 1910 through 1993', *Int. J. Climatol.*, **16**, 1005–1018.

Hughes, M.G., Frei, A. and Robinson, D.A. 1996. 'Historical analysis of North American snow cover extent: merging satellite and station derived snow cover observations', *Proceedings of the 1996 Eastern Snow Conference*, Williamsburg, VA, pp. 21–32.

Kukla, G. and Robinson, D.A. 1981. 'Accuracy of operational snow and ice charts', *1981 IEEE International Geoscience and Remote Sensing Symposium Digest*, pp. 974–987.

Livezey, R.E. 1995. 'The evaluation of forecasts', in Von Storch, H. and Navarra, A. (eds), *Analysis of Climate Variability: Applications of Statistical Techniques*, Springer-Verlag, Berlin, pp. 194–195.

Michaelson, J. 1987. 'Cross-validation in statistical climate forecast models', *J. Clim. Appl. Meteor.*, **26**, 1589–1600.

Pennell, W.T., Barnett, T.P., Hasselmann, K., Holland, W.R., Karl, T.R., North, G.R., MacCracken, M.C., Moss, M.E., Pearman, G., Rasmusson, E.M., Santer, B.D., Smith, W.K, von Storch, H., Switzer, P. and Zweirs, F. 1993. 'The detection of anthropogenic climate change', *Proceedings of the Fourth Symposium on Global Change Studies*, Anaheim, CA, American Meteorological Society, pp. 21–27.

Robinson, D.A. 1993a. 'Global snow cover monitoring: an update', *Bull Am Meteorol Soc.*, **74**(9), 1689–1696.

Robinson, D.A. 1993b. 'Historical daily climatic data for the United States', *Proceedings of the Eighth Conference on Applied Climatology*, Anaheim, CA, American Meteorological Society, pp. 264–269.

Robinson, D.A. 1999. 'Northern Hemisphere snow extent during the satellite era', *Fifth Conference on Polar Meteorology and Oceanography*, Dallas, TX, American Meteorological Society, pp. 255–258, preprints.

Smith, T.M., Reynolds, R.W., Livezey, R.E. and Stokes, D.C. 1996. 'Reconstruction of historical sea surface temperatures using empirical orthogonal functions', *J. Climate*, **V9**(6), 1403–1420.

Wiesnet, D.R., Ropelewski, C.F., Kukla, G.J. and Robinson, D.A. 1987. 'A discussion of the accuracy of NOAA satellite-derived global seasonal snow cover measurements', *Large Scale Effects of Seasonal Snow Cover*, *Proceedings of the Vancouver Symposium*, LAHS Publ. no. 166, pp. 291–304.

Ye, H., Cho, H.R. and Gustafson, P.E. 1998. 'Changes in Russian winter snow accumulation during 1936–83 and its spatial patterns', *J. Climate*, **V11**(5), 856–863.