



## Spatial variation in the influence of the North Atlantic Oscillation on precipitation across Greenland

Catherine A. Calder,<sup>1</sup> Peter F. Craigmile,<sup>1</sup> and Ellen Mosley-Thompson<sup>2,3</sup>

Received 27 July 2007; revised 9 November 2007; accepted 26 December 2007; published 29 March 2008.

[1] Ice core-derived accumulation records from Greenland have been proposed as proxies for North Atlantic Oscillation (NAO) reconstruction. In a series of single-site analyses, accumulation records from ice cores drilled in western Greenland were found to exhibit the strongest linear association with NAO. In this paper, we expand on these findings by proposing a spatiotemporal statistical model to explore further the relationship between NAO and the accumulation records from 35 firn and ice cores drilled in western and southern Greenland. In particular, we propose a temporal extension of the Bayesian spatially varying coefficient regression model, which is fit using a Markov chain Monte Carlo algorithm. This model readily accommodates the irregular features of the data (i.e., variation in record lengths and irregular spacing among ice core locations) and the serial dependence within individual records. Using our statistical model, we are able to exploit the spatial dependence structure of the derived accumulation-NAO relationship to explore the regional patterns in the strength of this relationship. Our findings support previous work that identified a region in western Greenland where derived accumulation is most correlated with NAO. However, we also identify a region further inland to the east and south where the estimated strength of the linear accumulation-NAO relationship is weaker, but more certain, than in the previously identified region. Thus, our findings can be used to guide decisions regarding where to locate future drilling efforts in Greenland by weighing the trade-off between the potential strength of the accumulation-NAO relationship and its uncertainty.

**Citation:** Calder, C. A., P. F. Craigmile, and E. Mosley-Thompson (2008), Spatial variation in the influence of the North Atlantic Oscillation on precipitation across Greenland, *J. Geophys. Res.*, *113*, D06112, doi:10.1029/2007JD009227.

### 1. Introduction

[2] In this paper, we propose a novel Bayesian hierarchical statistical model for exploring spatial variation in the relationship between the North Atlantic Oscillation (NAO), an important climatic phenomenon discussed below, and net annual accumulation across Greenland. This approach allows us to synthesize a large number of ice core-derived accumulation records in order to identify regional patterns in the strength of the accumulation-NAO relationship. As a result, inferences derived from our statistical model can be helpful when selecting specific ice core records to be used for NAO reconstruction and sites for future ice core drilling efforts.

[3] NAO refers to temporal changes in the atmospheric pressure differences between the region of low sea level pressure (SLP) near Iceland and the region of higher SLP in the subtropics. Changes in the strength of this meridional

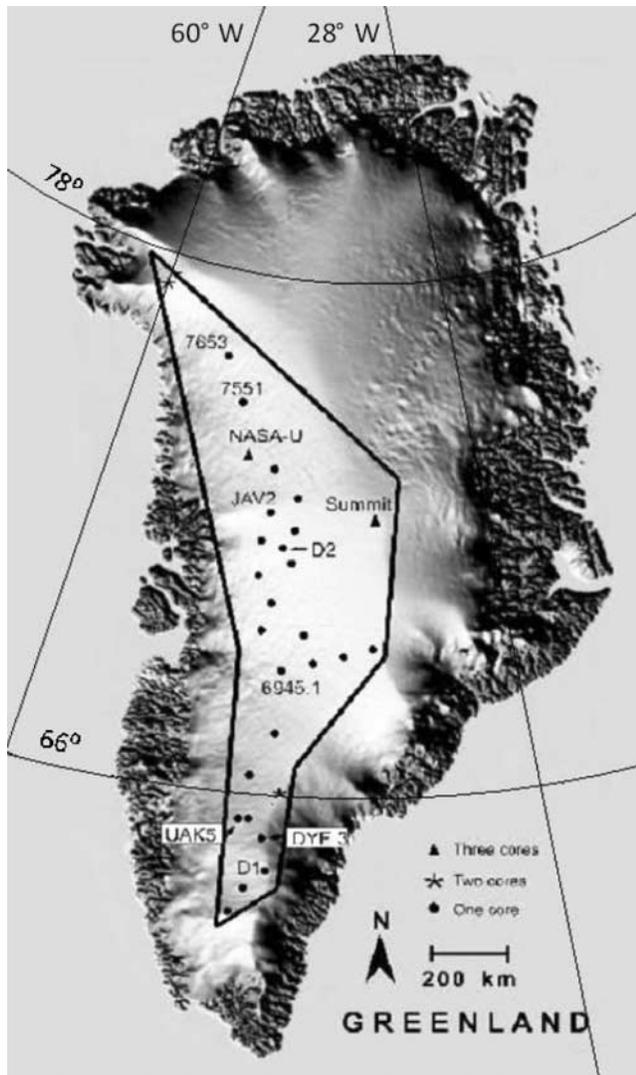
pressure dipole affect the zonal winds across the North Atlantic Ocean and the effects are felt throughout Europe and parts of Greenland [Hurrell, 1995]. NAO variability is also linked to changes in storm tracks [Rogers and van Loon, 1979; Hurrell et al., 2003], cyclone activity [Rogers, 1990; Serreze et al., 1997], precipitation patterns [Bromwich et al., 1999] and sea ice extent [Deser et al., 2000]. As a result, knowledge of the historical behavior of this system provides a context for understanding the implications of current and future climatic trends. NAO is characterized by an index derived as the difference in the mean monthly surface pressure anomalies between Iceland and the Azores. Various meteorological records have been used to construct NAO indexes [Rogers, 1984; Hurrell, 1995; Jones et al., 1997; Vinther et al., 2003a]. In addition, there have been several attempts to extend NAO indexes further into the past by substituting various climate proxies (e.g., tree rings, corals, ice cores) for the unavailable meteorological records [Appenzeller et al., 1998a, 1998b; Cook, 2003, and references therein; Vinther et al., 2003b; Mosley-Thompson et al., 2005, and references therein].

[4] Since precipitation over Greenland is modulated by the NAO, ice core-derived accumulation records from Greenland have been proposed as potential proxies for use in NAO reconstructions. Such records are available for a number of locations throughout Greenland. Some of the

<sup>1</sup>Department of Statistics, Ohio State University, Columbus, Ohio, USA.

<sup>2</sup>Byrd Polar Research Center, Ohio State University, Columbus, Ohio, USA.

<sup>3</sup>Also at Department of Geography, Ohio State University, Columbus, Ohio, USA.



**Figure 1.** Map of Greenland showing the sites where the cores used in this study were drilled. The Summit region includes three different cores (T2, T5 and GISP2). The black line delineates our study region.

older records (e.g., from Dye 3, see Figure 1) or records derived from cores located in central Greenland have been used for NAO reconstruction, with limited success. On the basis of an analysis of precipitation predicted using the ECMWF ERA-15 reanalysis (<http://www.ecmwf.int/research/era/Project/index.html>), [Appenzeller *et al.*, 1998a] demonstrated that the strongest correlation (negative) between precipitation and NAO is in the region west of central Greenland. Subsequent work [Appenzeller *et al.*, 1998b] found that the ice core-derived accumulation record from a PARCA (Program for Arctic Regional Climate Assessment) core located in western Greenland (NASA-U, see Figure 1) is more highly correlated with NAO than a composite accumulation record constructed from five cores drilled in the central Greenland (summit) region. Expanding on this finding, [Mosley-Thompson *et al.*, 2005] used five additional PARCA cores located across western Greenland to further explore the spatial and temporal variation in the correlation

between derived annual accumulation and NAO. Their analysis demonstrated that the relationship between NAO and accumulation has changed over time, especially as recorded at higher-latitude sites. However, despite these changes, the correlation between NAO and precipitation consistently remains strongest for ice cores drilled in western Greenland.

[5] A limitation of previous attempts to characterize the spatial variation in the relationship between NAO and precipitation over Greenland is that regional patterns were assessed on a site-by-site basis by examining inferences derived from single-core analyses. In these studies, the correlation between NAO and accumulation was computed at each site and then the variation in these single-site correlations over space was interpreted. The fact that NAO and accumulation were correlated at each site does not easily permit formal statistical inferences to be made about spatial patterns in the accumulation-NAO relationship. For example, it is not possible to assess whether the correlation between NAO and accumulation a region, as opposed to at one site, is statistically significant. Another limitation of this approach is that shorter ice cores must be excluded from the analysis since, taken separately, they contain little information. However, given the large number of shorter records available, it is worthwhile to explore whether they can be used in conjunction with the longer records to provide a more detailed description of the spatial patterns in the accumulation-NAO relationship.

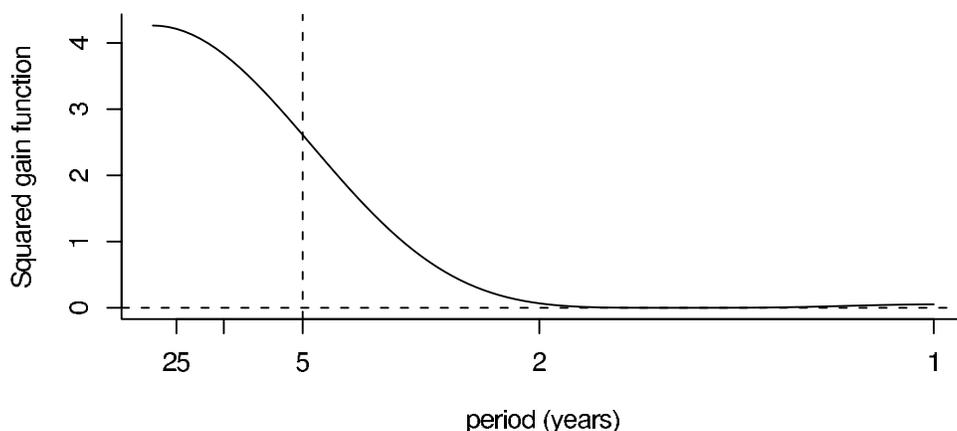
[6] Here, we explore the spatially varying relationship between ice core-derived net annual accumulation and NAO using a Bayesian hierarchical model. Our statistical model provides a mechanism for synthesizing the various accumulation records recovered across Greenland in a manner that accommodates the spatial dependence inherent in the accumulation process. Inferences regarding the strength of the linear association between NAO and accumulation at each location are derived using all cores collected at, or near the location of interest. In addition, our Bayesian model readily provides information about the spatial variation in the uncertainty in the accumulation-NAO relationship arising from the differing spatial and temporal coverage of the available ice core records.

[7] We describe the data used in our analysis in section 2. Then, in section 3, we develop a Bayesian hierarchical regression model for exploring the spatial variation in annual accumulation-NAO relationship and present summaries of our inferences in section 4. We conclude with a discussion of our results and propose directions for future research in section 5.

## 2. Data

[8] In our study, we use an annual (January to December) NAO index [Hurrell, 1995], that was obtained from <http://www.cgd.ucar.edu/cas/jhurrell/indices.data.html#naostatann>. This index dates back to 1865 and was constructed from normalized sea level pressures (SLP) between Ponta Delgada, Azores and Stykkisholmur/Reykjavik, Iceland.

[9] In addition, we use records of net annual mass accumulation (henceforth annual accumulation) derived from 35 firn and ice cores (henceforth cores), most of which were collected in western and southern Greenland



**Figure 2.** Squared gain function associated with the 5-point triangular filter with coefficients  $[1/\sqrt{19}, 2/\sqrt{19}, 3/\sqrt{19}, 2/\sqrt{19}, 1/\sqrt{19}]$ . The time periods for which the curve is highest are the periods that contribute the most to the filtered sequence. Thus, this filter emphasizes periods longer than 5 years (periods to the left of the vertical dotted line).

(Figure 1) by PARCA. The timescale for each core was established using seasonal variations in both chemical ( $\delta^{18}\text{O}$ ,  $\text{NO}_3^-$ ,  $\text{H}_2\text{O}_2$ ) and physical (insoluble dust) constituents, coupled with known volcanic horizons (see *Mosley-Thompson et al.* [2001, 2003] for details). For 32 of the 35 cores, the annual layer thicknesses were determined using the successive winter minima in both the insoluble dust concentration (dust) and the oxygen isotopic ratio ( $\delta^{18}\text{O}$ ). This resulted in two annual accumulation records for each core; each pair was then averaged to produce the annual accumulation record for each core used in this study. For the remaining three cores, only dust-derived (Summit Site T2 and T5 cores, see Figure 1) or  $\delta^{18}\text{O}$ -derived (Summit Site GISP2 core) annual layer thicknesses are available. The annual layers were converted to water equivalent (w.e.) using the density measurements for each core. The latitude/longitude coordinates for each of the core locations was converted to Universal Transverse Mercator (UTM) coordinates (Zone 26). This projection implies that the units for Euclidean distances between pairs of locations, which are needed in our statistical model, are kilometers.

[10] The variation in annual accumulation records derived from the ice cores results from a combination of processes operating at different temporal scales. For example, variations on shorter temporal scales may result from processes such as isotropic diffusion in the firm and drifting (erosion and redeposition), while variation at longer scales are more likely to reflect climatological processes. To isolate the longer-scale temporal variations, we follow the approach taken by *Appenzeller et al.* [1998a] (and subsequently *Mosley-Thompson et al.* [2005]). We first detrended each accumulation record by fitting a simple linear regression model (using ordinary least squares). The residual time series were then standardized to have a mean of zero and a variance of one. Given our interest in the variations (but not trends) over longer scales, a 5-point triangular filter with coefficients  $[1/\sqrt{19}, 2/\sqrt{19}, 3/\sqrt{19}, 2/\sqrt{19}, 1/\sqrt{19}]$  was applied to the standardized residual annual accumulation time series. (Filtering is the standard approach to analyze time series over different timescales [e.g., *Madden*, 1986]. If  $A_{i,j,t}$  is the net annual accumulation at location  $i$ , core  $j$ , at

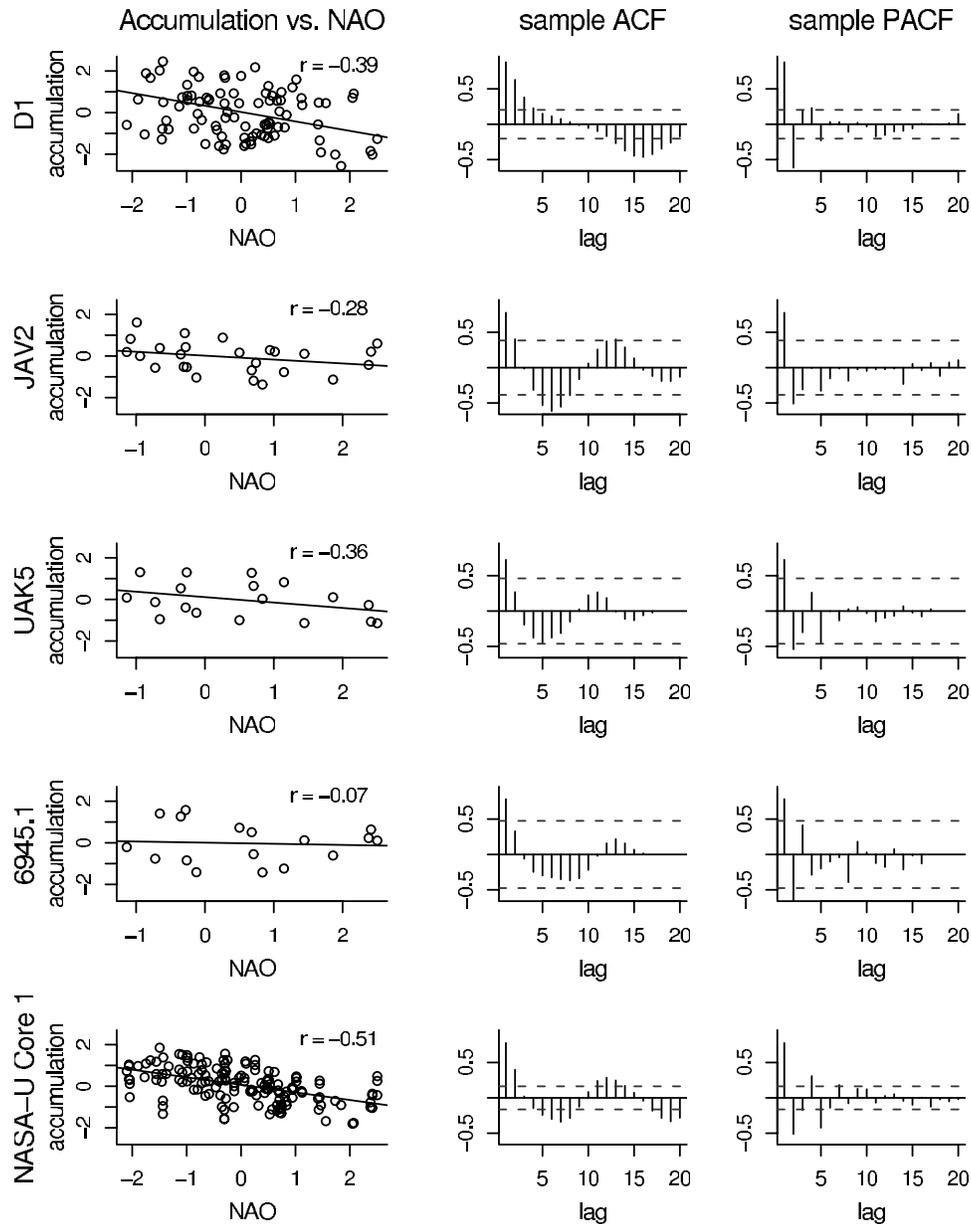
time  $t$  before filtering, then the filtered annual accumulation,  $A_{i,j,t}$  (used in our statistical analysis described in section 3) is defined to be

$$A_{i,j,t} = \frac{1}{\sqrt{19}} [\tilde{A}_{i,j,t-2} + 2\tilde{A}_{i,j,t-1} + 3\tilde{A}_{i,j,t} + 2\tilde{A}_{i,j,t+1} + \tilde{A}_{i,j,t+2}].$$

The effect of this filter on the time series can be explored by examining the associated squared gain function, which is defined as the modulus squared of the Fourier transform of the filter coefficients. Figure 2 displays the squared gain function associated with the 5-point triangular filter and reveals that it down-weights the high-frequency components in the signal (periods of less than 5 years), while emphasizing the decadal features. We carried out the same procedure (detrending, standardizing, and filtering with the 5-point filter) to the NAO index. As a result, our statistical analysis is designed to explore the relationship between NAO and annual accumulation at the decadal scale. We note that another way to isolate the decadal patterns is to directly model 10-year averages. However, a key disadvantage with this approach is that we lose a substantial amount of information by subsampling the data, especially in terms of the fidelity of the observed long-range signature.

[11] Filtering may introduce additional serial correlation, or autocorrelation, into the time series. The type of serial dependence introduced is dictated by the form of the filter. To illustrate this phenomenon, consider a time series consisting of a white noise, or uncorrelated, process. According to the linear-time-invariant filtering result for stationary time series [see *Percival and Walden*, 1993, chapter 5], after applying the 5-point triangular filter, the resulting process is a moving average process of order four, which is more correlated than the original unfiltered signal. Thus, in order to account for autocorrelation in our data, which may in part be due to the initial filtering, we include an underlying autoregressive process of order one in our statistical model, which is described in the next section.

[12] Before we introduce our statistical model for exploring the spatial variation in the strength of the relationship between the standardized NAO time series and the various



**Figure 3.** (left) Scatterplots of the filtered and standardized accumulation (in water equivalent) versus the NAO series for five cores in the study region. The solid line denotes the ordinary least squares line fit to each pair of time series. The sample correlation coefficient,  $r$ , is also provided. (middle and right) Sample ACF and PACF for the residuals from the ordinary least squares fit for each of the five cores.

standardized and filtered accumulation time series in section 3, we provide some graphical motivation for our choice of statistical model. Figure 3 (left) contains scatterplots of NAO versus accumulation for five ice cores (D1, JAV2, UAK5, 6945.1, and NASA-U Core 1) of different lengths. The ordinary least squares regression line has been added to each scatterplot to illustrate the linear relationship between NAO and accumulation. While there does appear to be some evidence of a linear relationship for all cores (sites), the strength of association, as measured by the correlation coefficient  $r$ , varies by the core (site). The two right hand panels display the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) for the residuals from the ordinary least squares regression line

for each core. In our statistical analysis, we model the residuals as an first-order autoregressive (AR(1)) process plus measurement error, which is equivalent to the residuals being a autoregressive moving average (ARMA(1,1)) process [Brockwell and Davis, 2002, exercise 2.9]. Note that the sample ACF is significantly different from zero at lag one and then decays like a sinusoid and the sample PACF is significant at lag one and then decays sinusoidally for longer lags. These patterns indicate that the AR(1) error plus measurement error model choice is a reasonable model for capturing the temporal dependence in the residual time series (see Brockwell and Davis [2002, chapter 5] for further discussion of model selection using the sample ACF and PACF). The lag one sample PACF is close to 0.7 for each

core (due in part to the fact that the same filter was used), which supports our assumption that the AR(1) partial autocorrelation parameter should be the same across sites.

### 3. Methods

[13] Recently, statistical models that can be broadly classified as spatially varying coefficient (SVC) models have been proposed as tools for exploring spatial variation in the relationship between explanatory variables and a response variable. One popular type of SVC model is the geographically weighted regression (GWR) model, which was proposed by *Fotheringham et al.* [2002]. A GWR model is constructed by fitting a series of weighted linear regression models associated with each location where the variables are observed. For a particular location, or focal location, the corresponding weights for the regression model are calculated using a kernel function that weights each of the other observations on the basis of its distance to the focal location. As a result of this weighting scheme, observations at locations that are close to the focal location contribute more in assessing the relationship between the explanatory variables and the response.

[14] Our analysis is based on a Bayesian alternative to GWR proposed by *Gelfand et al.* [2003]. This model is designed to explore the spatial variation in the relationship between explanatory and response variables that are spatially referenced. We build on this modeling framework by extending the standard Bayesian SVC model to accommodate temporal dependence in the process. Our approach allows us to account for the serial dependence in the accumulation and NAO records and is also amenable to the lack of consistency (length and spacing) among the ice core records.

[15] Following *Gelfand et al.* [2003], our extended SVC model is developed within the Bayesian setting. While the distinction between the Bayesian and classical, or frequentist, statistical paradigm is primarily philosophical in nature, the Bayesian approach has become increasingly popular because of its flexibility in specifying and fitting sophisticated, and scientifically driven, statistical models [e.g., *Wikle et al.*, 1998; *Berliner*, 2003]. Perhaps the most appealing feature of this approach is that inferences on unknown parameters reflect all sources of uncertainty in the model. As a result, our conclusions about the spatial variation in the influence of NAO on precipitation across Greenland are accompanied by spatially varying uncertainty values. These values reflect the various sources of uncertainty in the model due to the irregular spacing and differing lengths of the ice core records.

[16] Before introducing our extended Bayesian SVC model, we briefly review Bayesian inference. In the Bayesian paradigm, model parameters,  $\theta$  say, are considered to be unknown random quantities. Statistical inference proceeds by updating the prior distribution on the unknown parameters,  $p(\theta)$ , using the information in the observed data,  $Y$  say, as measured by the likelihood function,  $\mathcal{L}(Y|\theta)$ , to yield the posterior distribution of the parameters,  $p(\theta|Y)$ . This updating is performed in a coherent probabilistic manner using Bayes' Theorem:

$$p(\theta|Y) = \frac{\mathcal{L}(Y|\theta)p(\theta)}{p(Y)}.$$

[17] The denominator,  $p(Y)$ , in the above expression is equal to the integral of the numerator with respect to  $\theta$ . Typically, this integral cannot be evaluated in closed form, and thus simulation-based algorithms, such as those discussed in section 4, are used to approximate the posterior distribution of the unknown parameters.

[18] Returning to our extended Bayesian SVC model, we use the following notation:

- [19] Sets/indexes
- $\mathcal{D}$  the study region (see Figure 1);
  - $n_i$  number of ice core records available at location  $s_i \in \mathcal{S}$ ;
  - $\mathcal{S}$  set of  $m$  unique ice core locations, where  $\mathcal{S} \subset \mathcal{D}$ ;
  - $T_i^{\max}$  latest time when a derived annual accumulation value is available for any of the  $j$  ice cores at location  $s_i \in \mathcal{S}$ , i.e.,  $T_i^{\max} = \max(T_{i,1}^{\max}, \dots, T_{i,n_i}^{\max})$ ;
  - $T_{i,j}^{\max}$  latest time when a derived annual accumulation value is available for the  $j$ th ice core at location  $s_i \in \mathcal{S}$ ;
  - $T_i^{\min}$  earliest time when a derived annual accumulation value is available for any of the  $j$  ice cores at location  $s_i \in \mathcal{S}$ , i.e.,  $T_i^{\min} = \min(T_{i,1}^{\min}, \dots, T_{i,n_i}^{\min})$ ;
  - $T_{i,j}^{\min}$  earliest time when a derived annual accumulation value is available for the  $j$ th ice core at location  $s_i \in \mathcal{S}$ .

- [20] Data
- $A_{i,j,t}$  net annual accumulation (the smoothed record derived from annual layers identified using either dust or  $\delta^{18}\text{O}$ , or the average of both reconstructed histories) at time  $t$  as determined by the  $j$ th ice core at location  $s_i$ ;
  - $N_t$  value of the annual NAO index at time  $t$ .

- [21] Parameters
- $\beta$  spatial process representing the association between NAO and net annual accumulation;
  - $\gamma_t$  space-time process representing the net annual accumulation not explained by *NAO*;
  - $\lambda$  spatial dependence parameter of the  $\beta$  process;
  - $\mu$  mean of the  $\beta$  process;
  - $\sigma^2$  measurement error variance;
  - $\tau^2$  variance of the  $\beta$  process;
  - $\phi$  autocorrelation parameter of the time-varying  $\gamma$  process;
  - $\omega^2$  innovation variance of the time-varying  $\gamma$  process.

[22] We specify our Bayesian SVC model in a hierarchical manner, through a series of conditional distributions. This approach allows us to build a sophisticated statistical model by linking together various submodels. Throughout, we use the following notation that is frequently used in specifying Bayesian hierarchical models. For the random quantities  $X_1$  and  $X_2$ ,  $[X_1]$  denotes the distribution of  $X_1$ ,  $[X_1|X_2]$  denotes the conditional distribution of  $X_1$  given  $X_2$ , and  $[X_1, X_2]$  denotes the joint distribution of  $X_1$  and  $X_2$ .

[23] The first level of the hierarchy introduces a model for the accumulation data over space and time. Rather than *directly* specifying the joint distribution of the accumulation data, collectively denoted by  $\mathcal{A}$ , we model accumulation conditionally. At this level of the model, we assume that the spatial process  $\beta$ , the space-time process  $\gamma$ , and the error variance  $\sigma^2$  parameters are known. Then, given these parameters, the accumulation records are taken to be independent across space and time. In the subsequent levels of the hierarchy discussed below, we introduce models for  $\beta$ ,  $\gamma$

and  $\sigma^2$ , which indirectly imply a more sophisticated space-time dependence structure for the accumulation records. Using this conditioning approach, we assume that

$$[\mathcal{A}|N, \beta, \gamma, \sigma^2] = \prod_{i=1}^m \prod_{j=1}^{n_i} \prod_{t=T_i^{\min}}^{T_i^{\max}} [A_{i,j,t}|N_t, \beta(\mathbf{s}_i), \gamma_t(\mathbf{s}_i), \sigma^2], \quad (1)$$

where  $N$  denotes the NAO index. Each of the distributions on the right-hand side of (1) is assumed to be normal with mean  $\beta(\mathbf{s}_i)N_t + \gamma_t(\mathbf{s}_i)$  and variance  $\sigma^2$ ; we write

$$[A_{i,j,t}|\beta, \gamma, \sigma^2] = \mathbb{N}(\beta(\mathbf{s}_i)N_t + \gamma_t(\mathbf{s}_i), \sigma^2). \quad (2)$$

Together (1) and (2) determine the likelihood of the observed net annual accumulation records.

[24] In the second stage of the hierarchy, we specify models for  $\beta$ , the parameter capturing the spatial variation between annual accumulation and NAO, and  $\gamma$ , the parameter accounting for the residual spatiotemporal variation in annual accumulation not explained by NAO. Since these processes are themselves unobserved components of the model, these models can be thought of as prior distributions within the Bayesian paradigm. We assume that the spatially varying coefficient process,  $\beta$ , is Gaussian with a constant mean and a spatially dependent covariance structure. The latter assumption implies that the values of the  $\beta$  process at locations that are closer together in space are more highly correlated than the values of the process that are located further apart. In particular, we assume that the covariance between  $\beta(\mathbf{s})$  and  $\beta(\mathbf{s}')$ , for any  $\mathbf{s}, \mathbf{s}' \in \mathcal{D}$ , follows an exponential covariance model,

$$\text{cov}(\beta(\mathbf{s}), \beta(\mathbf{s}')) = \tau^2 \exp(-\|\mathbf{s} - \mathbf{s}'\|/\lambda), \quad (3)$$

where  $\|\mathbf{s} - \mathbf{s}'\|$  denotes the Euclidean distance (in km) between locations  $\mathbf{s}$  and  $\mathbf{s}'$ ,  $\tau^2$  is the unknown variance parameter, and  $\lambda$  is the spatial correlation (range) parameter. Thus, for the vector of values of the  $\beta$  process at the ice core record locations,  $\boldsymbol{\beta} = (\beta(\mathbf{s}_1), \dots, \beta(\mathbf{s}_m))^T$ ,

$$[\boldsymbol{\beta}|\mu, \tau^2, \lambda] = \text{MVN}(\mu\mathbf{1}, \mathbf{C}(\tau^2, \lambda)), \quad (4)$$

where  $\text{MVN}(\cdot, \cdot)$  denotes the multivariate normal distribution and  $\mathbf{1}$  represents a  $(m \times 1)$  vector of ones. The scalar parameter  $\mu$  captures the mean level of the  $\beta$  process, and the covariance matrix,  $\mathbf{C}(\tau^2, \lambda)$ , is determined by (3) and depends on the unknown spatial parameters,  $\tau^2$  and  $\lambda$ .

[25] The space-time process,  $\gamma$ , represents the residual net annual accumulation that is not explained by NAO. We assume *a priori* that this process is independent across space and follows a first-order autoregressive (AR(1)) model at each location  $\mathbf{s} \in \mathcal{D}$ :

$$[\gamma_t(\mathbf{s}_i)|\mathcal{H}_t(\mathbf{s}_i), \phi, \omega^2] = \begin{cases} \mathbb{N}(\phi\gamma_{t-1}(\mathbf{s}_i), \omega^2), & t = T_i^{\min} + 1, \dots, T_i^{\max}; \\ \mathbb{N}\left(0, \frac{\omega^2}{1 - \phi^2}\right), & t = T_i^{\min}. \end{cases}$$

[26] Here,  $\mathcal{H}_t(\mathbf{s}_i) = \{\gamma_r(\mathbf{s}_i): r = T_i^{\min}, \dots, t - 1\}$  denotes the historical values of the  $\gamma$  process at location  $\mathbf{s}_i$  that precede time  $t$ . In order to guarantee that the  $\gamma$  process is stationary at each location  $\mathbf{s}_i \in \mathcal{D}$ , we restrict the autocorrelation parameter  $\phi$  to lie between  $-1$  and  $1$ .

[27] We complete the specification of our Bayesian hierarchical model by assigning prior distributions for the remaining model parameters. For most of the parameters we use standard noninformative conjugate prior distributions:  $[\sigma^2] = \mathbb{IG}(0.001, 0.001)$ ,  $[\mu] = \mathbb{N}(0, 1000)$ ,  $[\tau^2] = \mathbb{IG}(0.001, 0.001)$ ,  $[\phi] = \mathbb{U}(-1, 1)$  and  $[\omega^2] = \mathbb{IG}(0.001, 0.001)$ . Here,  $\mathbb{IG}(a, b)$  denotes the inverse gamma distribution with shape parameter  $a$  and rate parameter  $b$ , and  $\mathbb{U}(a, b)$  denotes the continuous uniform distribution on the interval  $(a, b)$ . We assume  $[\lambda] = \mathbb{G}(5, 0.01)$ , where  $\mathbb{G}(a, b)$  denotes the gamma distribution with shape parameter  $a$  and rate parameter  $b$ . This assumption is designed to be fairly uninformative since to the best of our knowledge there is no relevant prior information available about the range of plausible values for this parameter. However, the prior distribution is not completely uninformative in the sense that it does induce some smoothness in the  $\beta$  process without being too restrictive. Finally, we note that the prior distributions for all parameters, except for  $[\lambda]$  and  $[\phi]$ , are conditionally conjugate. This property facilitates model fitting, as discussed in the next section.

## 4. Results

[28] Inferences about the unknown parameters in Bayesian models are based on the joint posterior distribution. In our extended Bayesian SVC model, the posterior distribution is

$$[\{\beta(\mathbf{s}_i); \mathbf{s}_i \in \mathcal{S}\}, \{\gamma_t(\mathbf{s}_i); \mathbf{s}_i \in \mathcal{S}, t = T_i^{\min}, \dots, T_i^{\max}\}, \sigma^2, \mu, \tau^2, \lambda, \phi, \omega^2 | \mathcal{A}, N]. \quad (5)$$

[29] Since this distribution cannot be derived in closed form, we approximate it using a simulation-based inference method known as Markov Chain Monte Carlo (MCMC). In an MCMC algorithm, a Markov Chain is constructed such that its stationary distribution is equal to the posterior distribution of interest. See, for example, *Gelman et al.* [1995] and *Chen et al.* [2000] for overviews of this methodology. In our analysis, we use a specific type of MCMC algorithm known as a Gibbs sampler in which we iteratively sample from the full conditional distributions of each of the unknown parameters (i.e., the distribution of an unknown parameter conditional on the values of the other unknown parameters and the data). The full conditional distributions for each of the unknown parameters in our model, except for the spatial dependence parameter  $\lambda$  and the autocorrelation parameter  $\phi$ , are available in closed form. In order to sample from the full conditional distributions of  $\lambda$  and  $\phi$ , we employ a random walk Metropolis step within our Gibbs sampler.

[30] In an MCMC algorithm, for a sample path of the Markov chain to constitute a random sample from the posterior distribution, the Markov chain must converge to its stationary distribution. As a result, the algorithm must run for a number of “burn-in” iterations before generating

**Table 1.** Bayesian SVC Model Results<sup>a</sup>

Model Component	Parameter	1900–	1925–
Measurement error	$\sigma^2$	0.190 (0.171, 0.211)	0.158 (0.140, 0.178)
$\beta$ process	$\mu$	-0.18 (-0.33, 0.02)	-0.18 (-0.33, 0.01)
$\beta$ process	$\lambda$	393 (121, 804)	432 (127, 905)
$\beta$ process	$\tau^2$	0.018 (0.001, 0.081)	0.013 (0.001, 0.049)
$\gamma$ process	$\phi$	0.78 (0.75, 0.81)	0.77 (0.73, 0.82)
$\gamma$ process	$\omega^2$	0.230 (0.201, 0.260)	0.248 (0.219, 0.281)

<sup>a</sup>Summaries of the posterior distributions (posterior means and 95 percent credible intervals) of the static model parameters for the extended Bayesian SVC model fitted to records dating back to 1900 and dating back to 1925.

posterior samples. In our model, the exact number of “burn-in” iterations is not known, and so we rely on several convergence diagnostic tools. One graphical method for assessing convergence is to look for trending in the trace plots of the sample paths of the various model parameters. In addition, to check that the chain was not stuck in a local mode, the algorithm can be run with different starting values. On the basis of these convergence diagnostics, we determined that “burn-in” was achieved after 5,000 iterations, and we then ran the algorithm for an additional 10,000 iterations. Since MCMC algorithms by construction produce dependent samples from the posterior, we store every tenth iteration to reduce the autocorrelation in posterior samples. After this thinning of our Markov chain, we had 1,000 nearly independent samples from the joint posterior distribution on which we base our inferences.

[31] Our model was fitted to all derived accumulation records dating back to 1900 A.D. We refer to this primary analysis as the post-1900 analysis. Given the large number of model parameters, we present only marginal distributional summaries of several key parameters. The third column in Table 1 lists the sample mean of the posterior samples of the static model parameters in the post-1900 analysis. In addition, estimated 95 percent credible intervals for each of these parameters are included. These intervals are approximated by the 2.5 and 97.5 percentiles of the posterior samples and provide information about the uncertainty in our knowledge of the parameters after observing the data. By definition, the probability that a parameter is in a 95 percent credible interval is 0.95, which implies that the wider the interval, the higher the uncertainty.

[32] Since our primary interest is assessing the influence of NAO on annual accumulation we also include posterior summaries of the  $\beta$  process. Our MCMC algorithm provides samples from the posterior of  $\beta(\mathbf{s})$ , for all observed  $\mathbf{s} \in \mathcal{S}$ .

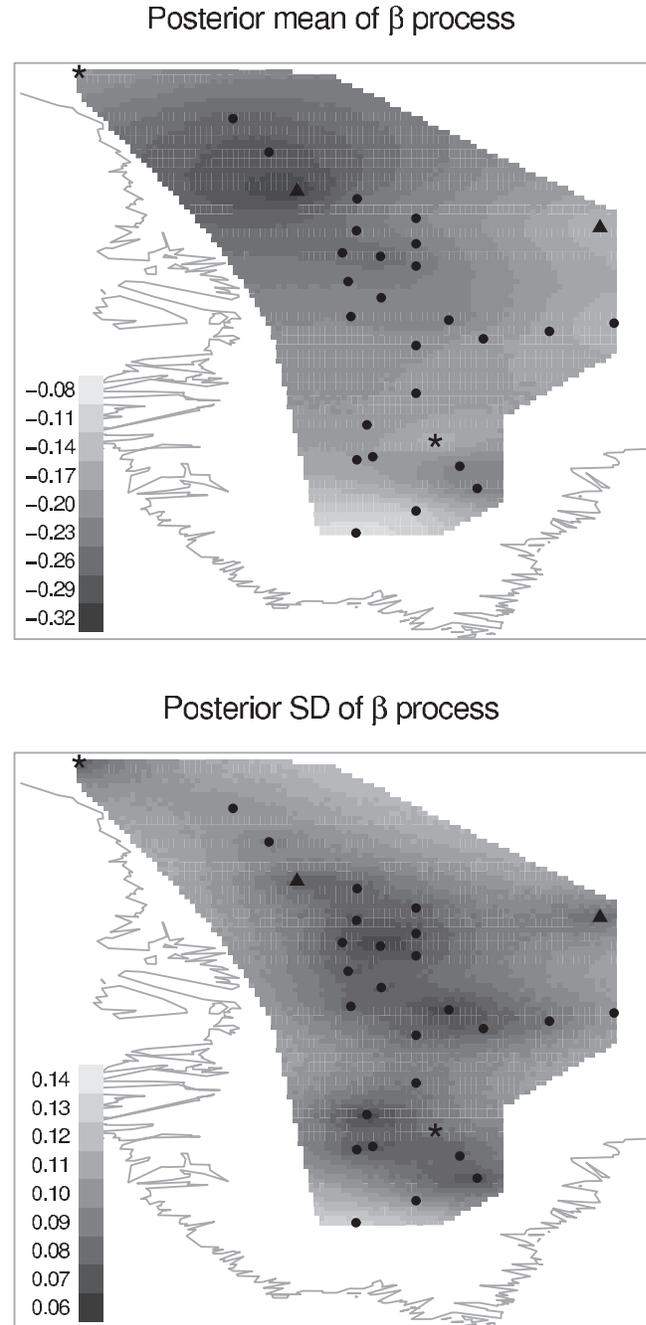
[33] To further explore the spatial relationship in the association between accumulation and NAO, we derive posterior summaries on a fine grid  $\{\mathbf{s}_1, \dots, \mathbf{s}_m\}$  covering the study region  $\mathcal{D}$ . The posterior predictive distribution of the  $\beta$  process on this grid is

$$p(\beta(\mathbf{s}_1), \dots, \beta(\mathbf{s}_m) | \mathbf{A}, \mathbf{N}) \propto \int p(\beta(\mathbf{s}_1), \dots, \beta(\mathbf{s}_m) | \mu, \tau^2, \lambda) dp(\mu, \tau^2, \lambda | \mathbf{A}, \mathbf{N}),$$

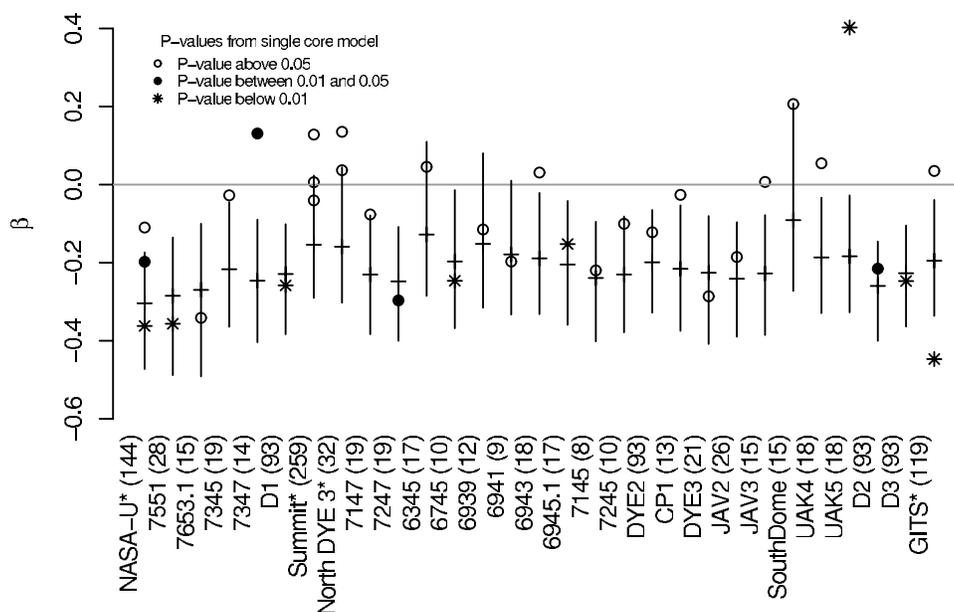
where  $p(\beta(\mathbf{s}_1), \dots, \beta(\mathbf{s}_m) | \mu, \tau^2, \lambda)$  is given by (4) and  $p(\mu, \tau^2, \tau^2, \lambda | \mathbf{A}, \mathbf{N})$  is the posterior distribution of  $\mu, \tau^2$ , and  $\lambda$ . We take a Monte Carlo approach to evaluating this integral.

First, we draw samples from the joint posterior of  $\mu, \tau^2$ , and  $\lambda$ . Then, conditional on these posterior samples, we draw values of the  $\beta$  process at locations  $\mathbf{s}_1, \dots, \mathbf{s}_m$  from the  $MV\mathcal{N}(\mu \mathbf{1}, \mathbf{C}(\tau^2, \lambda))$  distribution.

[34] Figure 4 provides summaries (means and standard deviations) of the posterior distribution of the gridded  $\beta$  process on a grid covering  $\mathcal{D}$ . Figure 4 reveals that the



**Figure 4.** Posterior (top) mean and (bottom) standard deviation of the  $\beta$  process on a grid covering the study region defined by the black lines on Figure 1. These inferences are based on fitting the extended Bayesian SVC model based on net accumulation records back to 1900 A.D. The black symbols indicate the core locations (see the legend in Figure 1) and are in the UTM coordinate system.



**Figure 5.** A summary of the  $\beta$  process from the fitted (using all available accumulation records back to 1900 A.D.) extended SVC model at each site. For each site, the horizontal dash denotes the posterior mean of  $\beta$  process, and the vertical line represents the corresponding 95 percent credible interval. The other symbols associated with each site indicate the maximum likelihood estimates of the strength of the accumulation-NAO relationship from individual ice core-specific ARMA(1, 1) time series regression models; the symbol indicates whether the relationship for each core is significantly different from zero (see section 4 for details). An asterisk next to the site label denotes sites with more than one core, and the number in parentheses is the total number of years in all the cores from that site.

relationship between NAO and net annual accumulation is strongest in the northwest part of  $\mathcal{D}$  in the region surrounding the NASA-U, 7551, and 7653 cores (see Figure 1), which is consistent with the conclusions of *Appenzeller et al.* [1998b] and *Mosley-Thompson et al.* [2005]. While the accumulation-NAO relationship is not as strong, the  $\beta$  process in the region to the south of NASA-U near D2 has a smaller posterior standard deviation (i.e., there is less uncertainty about the strength of the relationship).

[35] To demonstrate the utility of our Bayesian approach in synthesizing the available information about the accumulation-NAO relationship, we compare our findings to a set of comparable single-core analyses. Using the accumulation records for each core separately, we fitted an ARMA(1, 1) time series regression model of NAO on accumulation since the ARMA(1, 1) error structure is equivalent to an AR(1) model with measurement error. Figure 5 compares the posterior distributions of the extended SVC model  $\beta$  process to the maximum likelihood estimates of the parameter representing the strength of the accumulation-NAO relationship in the single-core analyses. For each core, the horizontal dash denotes the posterior mean and the vertical line represents the 95 percent credible interval of the  $\beta$  process at the core’s location, while the single-core estimates are denoted by one of three plotting symbols corresponding to the P value for a two-sided hypothesis test that the strength of the accumulation-NAO relationship is significantly different from zero. Figure 5 shows that the two analyses agree more consistently for locations with longer records from a single core or from multiple collocated cores (the numbers in parentheses next

to the site labels are the total number of accumulation years for all the cores at that site, and sites with an asterisk identify locations with multiple cores). Greater discrepancies between the inferences derived from the two approaches are observed for sites with fewer total records, especially in terms of the uncertainty levels. This phenomenon is particularly evident in the region to the south of NASA-U near D2. As we noted above, the strength of the accumulation-NAO relationship in this region is determined with a high level of certainty so the credible intervals are narrower. However, on the basis of the single-core analyses, the strength of the NAO-accumulation relationship in this region is almost never significantly different from zero at the 0.05 level since the cores in the region are relatively short. Thus, this comparison demonstrates the ability of the Bayesian extended SVC model to synthesize the information provided by the accumulation records in this region that, taken alone, are not as informative.

[36] Given the well-documented warming of the high northern latitudes in the 1920s, we follow *Mosley-Thompson et al.* [2005] and also fit our model to only the accumulation and NAO index values after 1925. The number of cores remain the same. The column on the right of Table 1 provides summaries of the posterior distributions of the static model parameters from this post-1925 analysis. These summaries, in addition to maps of the posterior mean and standard deviation of the  $\beta$  process (not provided for brevity), do not reveal any major differences in the accumulation-NAO relationship resulting from jointly analyzing all post-1900 records. Since our interest is in describing the spatial variation in the NAO influence on Greenland

accumulation, we do not further explore temporal variations in the relationship. Although future work might address this issue, the variation in the lengths of the accumulation records could prove to be problematic.

[37] In addition, we fit our model to the post-1900 data omitting the three core records that are based on a single seasonally varying parameter (dust or  $\delta^{18}\text{O}$ ; see the discussion in section 2). Again, omitting these cores from the analysis had very little effect on the results. As would be expected, the posterior standard deviation increased in the Summit region where the three cores were drilled (indicated by a triangle on Figure 1). However, the conclusions about the regions of Greenland where the NAO signature in the ice core-derived accumulation is strongest, as discussed above, do not change.

## 5. Discussion

[38] Our extended Bayesian SVC model provides a framework for examining the spatial variation in the NAO signature preserved in the Greenland ice sheet. Unlike previous analyses, our approach allows a formal assessment of the regional patterns in this relationship. In addition, it readily provides uncertainty statements about the strength of the accumulation-NAO relationship, which are shown to vary spatially.

[39] We draw two specific conclusions on the basis of our post-1900 analysis, which are discussed in section 4. First, the region where the relationship between NAO and ice core-derived accumulation is strongest is the dark area in the northwestern section of our study region. This result is consistent with those of *Appenzeller et al.* [1998b] and *Mosley-Thompson et al.* [2005], who both found that the accumulation histories from cores drilled at the NASA-U site in this region exhibited a strong correlation (negative) with NAO. However, our analysis reveals that the region where the uncertainty about the relationship between NAO and accumulation is lowest lies further to the southeast of NASA-U (bottom plot in Figure 4).

[40] It is possible that the reduced predictability of the accumulation-NAO relationship in the vicinity of NASA-U results, at least in part, from the dissimilarity among the three NASA-U records collected within a 2 km radius. The strength of the linear relationship among the cores, taken individually, and with NAO is variable. The Pearson correlations between each ice core-derived accumulation record and NAO are  $-0.524$ ,  $-0.286$ , and  $-0.162$  for cores 1, 2, and 3, respectively. (Statistical significance of these correlations is not provided because of the autocorrelation in the records.) In addition, the pairwise correlations among the individual records for their common 29-year period (1965–1993 AD) are variable (Cores 1 and 2: 0.706; Cores 1 and 3: 0.554; Cores 2 and 3: 0.854). The differences among the three accumulation histories are not the result of dating errors as the average accumulation (1965–1993 AD) is nearly identical (Cores 1, 2, and 3: 335, 331 and 323 mm w.e., respectively). The differences are real and largely result from surface processes that erode and redeposit surface snow such that a thicker annual layer at one site may be contemporaneous with a thinner layer at another site. The two cores northwest of NASA-U (sites 7653 and 7551) also have higher posterior standard deviations, con-

sistent with NASA-U, suggesting that the weak correlations among the NASA-U accumulation histories are not the primary cause of higher posterior standard deviations in that region.

[41] Although the longest record used in the study, NASA-U (Core 1, 93 years), exhibits a strong NAO signature in its accumulation record, the lack of consistency in this signature among the three records for their period of overlap implies that the accumulation-NAO relationship may be fairly uncertain (less robust) in the region. The lack of sufficiently long multiple cores at most sites makes comparable assessments of “local” consistency (robustness) difficult. It is reasonable to question whether or not the strong relationship between net accumulation and NAO at NASA-U (based on one longer core) is fortuitous, but the available data are insufficient to address this question and additional longer cores in the vicinity of NASA-U are needed to determine whether NASA-U is truly a sensitive location for NAO variability or whether the available record is anomalous.

[42] Our results suggest that a more consistent accumulation-derived NAO history is likely to be attained from cores drilled southeast of NASA-U and possibly in southern Greenland as well. A detailed investigation as to why this may be the case is beyond the scope of this paper, but linkages among the phase of the NAO, regional variability in sea ice extent and cyclone frequency offer a likely explanation. Observational evidence indicates that sea ice variability in the Arctic is modulated by large-scale atmospheric circulation patterns such as the Arctic Oscillation and the NAO [*Fang and Wallace*, 1994; *Rigor et al.*, 2002]. Further, Arctic sea ice extent is positively correlated with sea ice extent in the Greenland and Barents Sea region (east of Greenland) and negatively correlated with sea ice extent in the Labrador Sea–Baffin Bay region, south and west of Greenland creating the well known seesaw [*van Loon and Rogers*, 1978; *Walsh and Johnson*, 1979; *Wallace and Gutzler*, 1981; *Fang and Wallace*, 1994]. The extent and persistence of sea ice affects the regional climate by its influence on surface albedo and heat and moisture fluxes between the ocean and atmosphere which in turn modulate cyclone (storm) activity. *Deser et al.* [2000] used a principal component (PC) analysis to examine the relationship between winter sea ice concentrations and cyclone frequency for winters with high and low Arctic sea ice concentrations. The low minus high ice PC composite difference map reveals that west central Greenland (just south and east of the NASA-U region) experiences a maximum cyclone count as does a second center of action across southern Greenland [see *Deser et al.*, 2000, Figure 9c]. These are the regions of the ice sheet where our analysis yields lower posterior standard deviations (higher predictability) of the Beta process for the accumulation-NAO relationship. Thus this region appears more sensitive to temporal variations in the regional pressure distribution (NAO), sea ice extent and storm activity (cyclone frequency) and preserves a more consistent record of the accumulation-NAO relationship.

[43] Future efforts to use long ice core records from Greenland to extend the NAO history will need to consider our results carefully when selecting drill sites. Our analyses suggest that a new ice core in the NASA-U region (along the northwest side of Greenland) may contain an accumulation

history that is strongly correlated with NAO variability, but the higher levels of uncertainty here decrease the likelihood of success. Our model also suggests that a core further inland in the west central part of the ice sheet has a statistically better chance of capturing a reliable record of NAO variability.

[44] **Acknowledgments.** Peter F. Craigmile's research is supported by a National Science Foundation (NSF) grant, award 0604963. The PARCA cores were collected with support from NSF and NASA (NAGS-5032 and 6817) to Ellen Mosley-Thompson. This is contribution 1349 of the Byrd Polar Research Center. The authors are grateful for the anonymous reviewers' thoughtful comments and suggestions on the original version of the manuscript.

## References

- Appenzeller, C., J. Schwander, S. Sommer, and T. F. Stocker (1998a), The North Atlantic Oscillation and its imprint on precipitation and ice accumulation in Greenland, *Geophys. Res. Lett.*, *25*, 1939–1942.
- Appenzeller, C., T. F. Stocker, and M. Anklin (1998b), North Atlantic Oscillation dynamics recorded in Greenland ice cores, *Science*, *282*, 446–449.
- Berliner, L. M. (2003), Physical-statistical modeling in geophysics, *J. Geophys. Res.*, *108*(D24), 8776, doi:10.1029/2002JD002865.
- Brockwell, P. J., and R. A. Davis (2002), *Introduction to Time Series and Forecasting*, 2nd ed., Springer, New York.
- Bromwich, D. H., Q. Chen, Y. Li, and R. L. Cullather (1999), Precipitation over Greenland and its relation to the North Atlantic Oscillation, *J. Geophys. Res.*, *104*(D18), 22,103–22,116.
- Chen, M.-H., Q.-M. Shao, and J. G. Ibrahim (2000), *Monte Carlo Methods in Bayesian Computation*, Springer, New York.
- Cook, E. R. (2003), Multi-proxy reconstruction of the North Atlantic Oscillation (NAO) index: A critical review and a new well-verified winter NAO index reconstruction back to AD 1400, in *The North Atlantic Oscillation: Climate Significance and Environmental Impact*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., pp. 63–79, AGU, Washington, D. C.
- Deser, C., J. E. Walsh, and M. S. Timlin (2000), Arctic sea ice variability in the context of recent atmospheric circulation trends, *J. Clim.*, *13*, 617–633.
- Fang, Z., and J. M. Wallace (1994), Arctic sea ice variability on a timescale of weeks and its relation to atmospheric forcing, *J. Clim.*, *7*, 1897–1914.
- Fotheringham, A. S., C. Brunson, and M. Charlton (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley, Hoboken, N. J.
- Gelfand, A. E., H. Kim, C. F. Sirmans, and S. Banerjee (2003), Spatial modeling with spatially varying coefficient processes, *J. Am. Stat. Assoc.*, *98*, 387–396.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (1995), *Bayesian Data Analysis*, Chapman and Hall, London.
- Hurrell, J. W. (1995), Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, *269*, 676–679.
- Hurrell, J. W., Y. Kushnir, M. Visbeck, and G. Ottersen (2003), An overview of the North Atlantic Oscillation, in *The North Atlantic Oscillation: Climate Significance and Environmental Impact*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., pp. 1–35, AGU, Washington, D. C.
- Jones, P. D., T. Jonsson, and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observation from Gibraltar and south-west Iceland, *Int. J. Clim.*, *17*, 1433–1450.
- Madden, R. A. (1986), Seasonal variation of the 40–50 day oscillation in the tropics, *J. Atmos. Sci.*, *43*, 3138–3158.
- Mosley-Thompson, E., J. R. McConnell, R. C. Bales, Z. Li, P. Lin, K. Steffen, L. G. Thompson, R. Edwards, and D. Bathke (2001), Local to regional-scale variability of Greenland accumulation from PARCA cores, *J. Geophys. Res.*, *106*(D24), 839–851.
- Mosley-Thompson, E., T. A. Mashiota, and L. G. Thompson (2003), Ice core records of late Holocene volcanism: Current and future contributions from the Greenland PARCA cores, in *Volcanism and the Earth's Atmosphere*, *Geophys. Monogr. Ser.*, vol. 139, edited by A. Robock and C. Oppenheimer, pp. 153–164, AGU, Washington, D. C.
- Mosley-Thompson, E., C. R. Readinger, P. Craigmile, L. G. Thompson, and C. A. Calder (2005), Regional sensitivity of Greenland precipitation to NAO variability, *Geophys. Res. Lett.*, *32*, L24707, doi:10.1029/2005GL024776.
- Percival, D. B., and A. Walden (1993), *Spectral Analysis for Physical Applications*, Cambridge Univ. Press, Cambridge, U. K.
- Rigor, I. G., J. M. Wallace, and R. L. Colony (2002), Response of sea ice to the Arctic Oscillation, *J. Clim.*, *15*, 2648–2663.
- Rogers, J. C. (1984), The association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere, *Mon. Weather Rev.*, *112*, 1999–2015.
- Rogers, J. C. (1990), Patterns of low-frequency monthly sea level pressure variability (1899–1986) and associated wave cyclone frequencies, *J. Clim.*, *10*, 1635–1645.
- Rogers, J. C., and H. van Loon (1979), The seesaw in winter temperatures between Greenland and northern Europe. Part II: Some oceanic and atmospheric effects in middle and high latitudes, *Mon. Weather Rev.*, *107*, 509–519.
- Serreze, M. C., F. Carsey, R. G. Barry, and J. C. Rogers (1997), Icelandic low activity: Climatological features, linkages with NAO, and relationships with recent changes in the Northern Hemisphere circulation, *J. Clim.*, *10*, 453–464.
- van Loon, H., and J. C. Rogers (1978), The seesaw in winter temperatures between Greenland and northern Europe. Part 1: General description, *Mon. Weather Rev.*, *106*, 296–310.
- Vinther, B. M., K. K. Andersen, A. W. Hansen, T. Schmith, and P. D. Jones (2003a), Improving the Gibraltar/Reykjavik NAO index, *Geophys. Res. Lett.*, *30*(23), 2222, doi:10.1029/2003GL018220.
- Vinther, B. M., S. J. Johnsen, K. K. Anderson, H. B. Clausen, and A. W. Hansen (2003b), NAO signal recorded in stable isotopes of Greenland ice cores, *Geophys. Res. Lett.*, *30*(7), 1387, doi:10.1029/2002GL016193.
- Wallace, J. M., and D. S. Gutzler (1981), Teleconnection in the geopotential height field during the Northern Hemisphere winter, *Mon. Weather Rev.*, *109*, 784–812.
- Walsh, J. E., and C. M. Johnson (1979), Interannual atmospheric variability and associated fluctuations in Arctic sea ice extent, *J. Geophys. Res.*, *84*, 6915–6928.
- Wikle, C. K., L. M. Berliner, and N. Cressie (1998), Hierarchical Bayesian space-time models, *Environ. Ecol. Stat.*, *5*, 117–154.

C. A. Calder and P. F. Craigmile, Department of Statistics, Ohio State University, Columbus, OH 43210, USA.

E. Mosley-Thompson, Byrd Polar Research Center, Ohio State University, Columbus, OH 43210, USA.