S6. EXTREME NORTH AMERICA WINTER STORM SEASON OF 2013/14: ROLES OF RADIATIVE FORCING AND THE GLOBAL WARMING HIATUS

Xiaosong Yang, G. A. Vecchi, T. L. Delworth, K. Paffendorf, R. Gudgel, L. Jia, Seth D. Underwood, and F. Zeng


1. Model description
We used the Geophysical Fluid Dynamics Laboratory (GFDL) Forecast-oriented Low Ocean Resolution model (FLOR; Vecchi et al. 2014). FLOR comprises 50-km mesh atmosphere and land components, and 100-km mesh sea ice and ocean components. The atmosphere and land components of FLOR are taken from the Coupled Model version 2.5 (CM2.5; Delworth et al. 2012) developed at GFDL, whereas the ocean and sea ice components are based on the GFDL Coupled Model version 2.1 (CM2.1; Delworth et al. 2006; Wittenberg et al. 2006; Gnanadesikan et al. 2006).

Two versions of FLOR were used for the initialized and uninitialized climate simulations. One is the version B01, called FLOR_B01 (Vecchi et al. 2014). The other is the FLOR_B01 model with “flux adjustments” (Magnusson et al. 2013; Vecchi et al. 2014), which adjusts the model’s momentum, enthalpy, and freshwater fluxes from atmosphere to ocean and so brings the long-term climatology of sea surface temperature (SST) and surface wind stress closer to the observations. The flux-adjusted FLOR model is called FLOR_FA in the text.

2. Experimental settings

a. Retrospective seasonal forecasts. For each year in the period 1980–2014, we used the results of 12-member ensemble retrospective seasonal forecasts by FLOR (Vecchi et al. 2014; Jia et al. 2015). For each ensemble member, 12-month duration predictions were performed after initializing the model to observationally constrained conditions. The 12-member initial conditions for ocean and sea ice components were built through a coupled data assimilation system developed for CM2.1 using an ensemble Kalman filter (EnKF), whereas those for atmosphere and land components were built from a suite of SST-forced atmosphere-land-coupled simulations using the components in FLOR, which is called Phase-1 (P1) initialization scheme. In addition, we used the same ocean/ice initial conditions as P1, but generated initial conditions of atmosphere and land components by nudging atmosphere component to the MERRA reanalysis with the same SST-forcing as P1, which is called Phase-2 (P2) initialization scheme.

To compute the extreme probability for each December–February (DJF) between 1990 and 2014, we used 12-member ensemble forecasts initialized on the first day of December and November from FLOR_B01 and FLOR_FA with P1 and P2 schemes, thereby yielding 84 samples for each predicted year. Note that the hindcasts of FLOR_B01 P2 starting 1 November were not available.

b. 1860/1990-control simulations. We generated 2000-year control climate simulations using FLOR by prescribing radiative forcing and land-use conditions representative of the year 1860. In addition, we generated 500-year control climate simulations by prescrib-
ing conditions representative of the year 1990. For these experiments, we compute the probability of the 2013/14-like extreme event \( P(\text{ETSI of 1314}) \) using all simulated years. To elucidate multicentury variability, we compute \( P(\text{ETSI of 1314}) \) for each 100-year chunk. The bars in Fig. 6.2 of main text show the standard errors in the variability.

c. Large-ensemble historical forcing simulations. We conducted 35-member ensemble historical forcing simulations using FLOR_FA. Five- (thirty-) member ensemble simulations were conducted from 1861 (1941) to 2040 by prescribing historical anthropogenic forcing and aerosols up to 2005, and future levels based on the Radiative Concentration Pathways scenario 4.5 (RCP4.5) from 2006 to 2040. In the simulations, historical volcanic radiative forcing was also prescribed up to 2005; however, no volcanic forcing was prescribed after 2006. These historical simulations were not initialized by the coupled data assimilation. They are called ALLFORC in the main text. For every 10 years from 1950 to 2030, we collect a centered 20-year data \((35 \times 20 = 700 \text{ samples})\), then split the 700 samples into seven 100-chunk for calculating \( P(\text{ETSI of 1314}) \). The mean and error bars of \( P \) are computed from the seven chunks.

d. Global warming hiatus simulations. The global warming hiatus simulations were from Delworth et al. (2015). The simulations are identical to ALLFORC, except that the model calculated wind stress anomalies that the ocean feels are replaced over the tropical Pacific with stress anomalies derived from reanalysis. The replacement of wind stress anomalies occurs only over the tropical Pacific, and only affects the momentum flux. These simulations are called ALLFORC_STRESS, and cover the period 1979–2013. The experiments were repeated in two versions of FLOR model, namely FLOR_B01 and FLOR_FA, with 35, and 5 ensemble members respectively.

3. Reliability of the probability estimates by model

To assess the reliability of these probability estimates in the retrospective forecasts, we use quantile–quantile plots that compare the probability of exceedance in the predicted cumulative distribution function (CDF) for the ETSI that were observed in that year (verification exceedance probability) with the sorted ranking (normalized by total number of forecasts) of the verification exceedance probabilities. If the forecast CDFs were reliable, for a large enough sample size, the points are expected to lie on the diagonal. The CDFs of the observed value on the predicted probability are generally close to a uniform distribution for all three cases (Fig. S6.1a–c): the forecast CDFs appear reliable.

![Quantile-quantile plots](image-url)
4. Extreme event definition in uninitialized simulations
In the retrospective forecasts, the extreme values of ETSI are close to observed values due to the initialization. However, the extreme values of ETSI in uninitialized simulations have biases, so we calibrated the extreme values based on the empirical occurrence probability of the observed 2013/14 event in observations. We define the extreme events in these uninitialized models using the observed 2013/14 extreme occurrence probability values 3%, 95%, and 90% for the Pacific, mid-USA, and Canada respectively, during 1980–2010 in the 35-member historical forcing simulations. Note that Pacific is in the negative extreme category, we compute the extreme occurrence probability over Pacific using the following equation:

\[ P(x) = \frac{\text{Number of years with ETSI} \leq x}{\text{Total number of years}}. \]

5. Response to the tropical Pacific wind stress anomalies during global warming hiatus
We computed the mean response of ETS to the tropical Pacific wind stress anomalies by subtracting the ensemble mean ETS of ALLFORC from the ensemble mean ETS of ALLFORC_STRESS during the global warming hiatus period (2000–12) (Fig. S6.2). The spatial patterns of the response are quite similar, but the amplitudes of FLOR_B01 are generally larger than those of FLOR_FA. For the 35-member FLOR_FA, the extreme probabilities \( P(\text{ETSI of 1314}) \) of the three regions were computed for the DJF from five 7-member chunks (each chunk contains \( 7 \times 13 = 105 \) samples) during the global warming hiatus period and for ALLFORC_STRESS and ALLFORC experiments respectively, and then we computed the mean and standard errors of the estimated probability from the five chunks. For the 5-member FLOR_B01 experiments, only one estimated probability value was calculated due to the limited sample size.

To provide a measure of the attributable risk, we computed the fraction of attributable risk (FAR) between the ALLFORC and ALLFORC_STRESS as

\[ \text{FAR} = \frac{P(x | \text{ALLFORC_STRESS}) - P(x | \text{ALLFORC})}{P(x | \text{ALLFORC_STRESS})}. \]
6. Measures for extratropical storms

The extratropical storm index (ETSI) is defined as the seasonal standard deviation of filtered 6-hourly sea level pressure (SLP) field, and the filter is a 24-hour difference filter (Wallace et al. 1988). The ETSI can be written as follows:

\[
ETSI = \sqrt{\frac{1}{N-1} \sum_{t=1}^{N} [SLP(t + 24hr) - SLP(t)]^2},
\]

where \( N \) is the sample size of the December–February (DJF) for each year.

To test the sensitivity to the methods of defining ETSI, we applied the conventional 2–6-day bandpass filter to the observed SLP (Hoskins and Hodges 2002) and the filter is a 128-point Lanczos filter. Figure S6.3 shows climatological DJF ETSI estimated from the two methods. It is evident that the two methods are capable of capturing the well-known intensified ETS activities over North Pacific, North Atlantic, and the midlatitudes in Southern Hemisphere. The global geographic distribution of ETS is almost the same between the two methods.

We further compare the anomalous ETS of 2013/14 DJF over North America between the two methods (Fig. S6.4). The spatial patterns of the 2013/14 anomalies over North America are similar between the two, and the time series of ETS over the three regions of interest are highly correlated with each other with correlation coefficients of 0.94 for mid-USA, 0.94 for mid-Canada, and 0.92 for Pacific. In addition, the two methods agree on the extreme events of ETS over the three regions. Thus, the 24-hour-difference filter is a robust filter for identifying ETS.

**Fig. S6.4.** The 2013/14 DJF ETSI anomalies derived from the 6-hourly SLP of ERA-Interim data using (a) a 24-hour-difference filter and (b) a 2–6-day bandpass filter. The time series of DJF ETSI anomalies using a 24-hour-difference filter (blue) and a 2–6-day bandpass filter (red) for (c) mid-USA, (d) mid-Canada, and (e) Pacific. Units are hPa.
REFERENCES


