nature climate change

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCLIMATE2483

Influence of internal variability on Arctic sea-ice trends

Neil C. Swart *[1], John C. Fyfe[1], Ed Hawkins[2], Jennifer E. Kay[3], and Alexandra Jahn[4]

[1] *Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia V8W 2Y2, Canada.*
[2] *National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading RG6 6BB, UK.*
[3] *Dept. Atmospheric and Oceanic Sciences, University of Colorado at Boulder, Boulder, Colorado, USA.*
[4] *National Center for Atmospheric Research, Boulder, Colorado, USA.*

November 17, 2014

1 Data and methods

We use the observed Northern Hemisphere September mean sea-ice extent data from the National Snow and Ice Data Center[1]. For 31 CMIP5 models we calculated the sea-ice extent as the area-integral of all grid cells in the Northern Hemisphere with a sea-ice concentration of 15% or greater. The CMIP5 data comprised of the historical simulation from 1979 to 2005 and the RCP2.6, RCP4.5 and RCP8.5 experiments from 2006 to 2100. The number of realizations available for each CMIP5 model and experiment is listed in Table S1. We similarly calculated sea-ice extent from 30 realizations of the CESM1 Large Ensemble, which comprised of the historical experiment over 1979 to 2005 and RCP8.5 from 2006 to 2013.

*Neil.Swart@ec.gc.ca

DOI: 10.1038/NCLIMATE2483

Figure 1 and 2 use the historical-RCP4.5 experiments from CMIP5. In Figure 2 the CESM1 LE results use the historical-RCP8.5 experiments. Since the RCP scenarios are very similar over 2006 to 2013, the exact choice of scenario has little influence on the results over this period. For example, this can be seen in Figure 3 in the main text for CMIP5. The application of external forcing (greenhouse gas, aerosol, solar and volcanic) may differ slightly between different models and between models and observations. For example, the aerosol and solar forcing have differed between the RCP scenarios and the observations since 2005[2].

For Figure 4 only a subset of six GCMs were chosen because they were the only models to produce multiple realizations for all 3 RCPs considered, which also had at least one pentad overlapping with the observations between 1986 and 2005. They are indicated by a '*' in Table S1. The GISS models have multiple members for each scenario but they are not included because those members represent a perturbed physics ensemble, rather than a multi-realization ensemble like the other models (i.e. same physics, perturbed initial condition).

1.1 Probability calculation

Figure 2 shows linear trends in sea ice extent that were calculated for all 7 and 14 year periods between 1979 and 2013, yielding 29 and 22 trends respectively for each model realization and the observations. The probabilities, p, were calculated as the number of trends equal to or exceeding the given threshold (e.g. ≥ 0), divided by the total number of trends (i.e. 29 or 22). In Figure 3b probabilities were calculated for all 7 year trends in a rolling 21 year window. In all cases probabilities were calculated for each realization, and then averaged across the model ensemble. For the CMIP5 models, the ensemble mean probability was calculated as a weighted mean such that each model held a weight of 1. Thus, models with many realizations contributed equally to the mean as models with few realizations. Similarly, the ensemble mean sea-ice extent in Figure 1 represents a weighted mean. However, this weighting has only a small effect on the probabilities, and all conclusions remain the same when using unweighted averages. Probabilities for CESM1 LE were calculated in the same way, except that no weighting is applied. Furthermore probabilities were only calculated for trends where sea-ice extent exceed 1×10^6 km^2, to avoid "ice-free" conditions from biasing the results.

2 Historical versus future trends

The relationship between the minimum, maximum and mean 7 year trend over 1979 to 2013 with the future trend over 2013 to 2070 is shown in Figure S1. Neither the minimum, maximum nor the mean historical trends have a significant relationship with the future trend at the 5% level, as given by the p-valuef of the test of the null hypothesis that the slope of the regression line is equal to zero. The coefficient of determination, R^2, of the fit between historical and future trends also has values close to zero, indicating that historical short term trends have little relation to future long term trends.

3 Comparing observed and simulated trends

We follow Fyfe et al.[3] in testing whether observed and CMIP5 simulated trends are equal. We assume exchangeability between models, in which case a reasonable representation of the trends is:

$$M_{ij} = u^m + Eint_{ij} + Emod_i,\ 1, ..., N^m,\ 1, ..., N_i \text{ and} \tag{1}$$

$$O = u^o + Eint^o \tag{2}$$

where M_{ij} and O are trends calculated from single runs or the observations. u^m and u^o are the true, unknown, deterministic trends due to external forcing in the model and observations, and u^m is the component of the trend common to all models (in the limit as the collection of exchangeable models grows infinitely large). Thus u^m is essentially the model-mean trend, as we will see below. $Eint_{ij}$ and $Eint^o$are perturbations to M_{ij} and O respectively due to internal variability, and for the models this is different for each run. $Emod_i$ is the perturbation to M_{ij} that is introduced by model error in model i. We assume that these perturbations are exchangeable. Nm is the number of models, and Ni is the number of realizations for model i. An estimator of $u^o - u^m$ is $O - M_{..}$, where "." replacing the subscripts represents averaging over that subscript. The null hypothesis that the observed and simulated trends are equal is

$$H_0 : u^m = u^o \tag{3}$$

which can be tested using an empirical distribution that includes the sources of uncertainty in (1) and (2). Using a resampling approach to build the empirical distribution avoids making any distributional assumptions. The empirical distribution is constructed as follows:

DOI: 10.1038/NCLIMATE2483

a) Compute the observed trend, O; b) Select a sample of Nm models with replacement and for each selection, draw one run at random from that model's available ensemble of realizations, and then average over those Nm runs to obtain a version of $M_{..}$; c) Select, at random, a single model i from models with multiple simulations, and then select, at random, a single run j from that model's ensemble. Calculate the difference $M_{ij} - M_{i.}$ between the trend in that single run and the mean of the trends from that model's ensemble. This difference is an estimate of the deviation in the j-th trend for model i that is induced by internal variability. Since the model i ensemble is generally small, the deviations are smaller than would be representative of an infinitely large replication of runs for model i, and so to compensate for that loss of variance, multiply the difference $M_{ij} - M_{i.}$ by $[Ni/(Ni-1)]^{0.5}$. d) Calculate $a - b + c$, as computed in the steps above, and repeat many times to build a distribution for $a-b+c$. From this distribution we compute p-values for the test of the null hypothesis, H_0. Smaller p-values are stronger evidence against the null hypothesis. For the rational for this procedure see Fyfe et al.[3].

For CESM1 LE we proceed as above, except that there is only one model from which to resample. Thus b) is a random sample with replacement of N_i CESM1 LE runs and c) is as above but always selecting, at random, a single run j from CESM1 LE.

4 Dependence on mean state, variance, and long term trend

The CMIP5 ensemble mean September sea-ice extent climatology over the historical period is very close to observations, as is the historical ensemble mean variance (Fig. S2). However, there is a large spread across the model ensemble both in terms of climatology and variance, with some models having much higher or lower than observed values (also see Fig. S5). The CESM1 LE has ensemble mean values close to observations, and a small spread due only to internal variability.

Given the large spread amongst the CMIP5 models, we may ask: are the probabilities calculated for a pause in sea-ice extent loss sensitive to the model base state? In fact the probability of a 7-year period featuring a pause is not sensitive to the historical mean sea-ice extent or its variance in the CMIP5 ensemble (Fig. S3a, b). The probability of a pause is weakly though significantly correlated with the long-term background trend over 1979 to 2013 (Fig. S3c). As discussed in the main text, the observed and

simulated long term trends are not inconsistent with each other. Thus the weak dependence of a pause on the background trend should not bias the calculated probabilities.

Another way of seeing that the probabilities are nearly independent of the mean state is to consider the CESM1 LE. That model has a small range of climatologies and variances amongst its 30 members, but has a similar range of probabilities to CMIP5 (Fig. S3a, b). Finally, as shown below, a large part of the spread in CMIP5 7 year trends is driven by internal variability rather than inter-model spread (Fig. S4). Thus, while model mean states differ significantly, model mean short terms trends are similar. The fact that the probabilities are nearly independent of the mean state also means that our results are not sensitive to the selection of models.

5 Internal variability versus inter-model spread

5.1 Sea-ice extent trends

The spread of 7 year trends in CMIP5 is driven by internal variability and inter-model spread. To consider which of these is more important, we first compare the CMIP5 trend distribution with the CESM1 LE distribution, the latter of which only contains spread only due internal variability. The spread of 7 year trends in the CESM1 LE is the similar to that in CMIP5 (Figure 2a), with standard deviations of 1.2 and 1.4×10^6 km^2 dec. $^{-1}$ respectively. This suggests that the CMIP5 spread could plausibly arise in large part due to internal variability. For further insight we compare the distribution of 7 year trends amongst all CMIP5 models with more than 10 realizations (Fig. S4). The models have a similar mean trend, but within each model the spread of trends amongst the realizations is large. This confirms the notion that the distribution of 7 year trends is primarily driven by inter-realization spread (i.e. internal variability), rather than by inter-model spread.

Similarly, the spread of 14 year trends is 0.55 and 0.63×10^6 km^2 dec. $^{-1}$ in CESM1-LE and CMIP5 respectively. For 35 year trends the spread is 0.21 and 0.27×10^6 km^2 dec. $^{-1}$ in CESM1-LE and CMIP5 respectively. Thus, even for multi-decadal trends, it is plausible that a large fraction of the CMIP5 spread arises due to internal variability.

5.2 Mean sea-ice extent

The distribution of mean sea-ice extents across the CMIP5 ensemble behaves differently to sea-ice extent trends, particularly on multi-decadal timescales.

DOI: 10.1038/NCLIMATE2483

Over 1980 to 2010 there is a large range in the climatological sea-ice extents across the CMIP5 ensemble (Fig. S2), with a total range of 10 million km^2. In contrast, the range in climatological extents across the 30 CESM1-LE realizations is less than 1 million km^2. Quantified in terms of standard deviation, the spreads over this period are 0.2 and 2.6 million km^2 in CESM1-LE and CMIP5 respectively. This suggests that inter-model spread is large relative to inter-realization spread (internal variability) for mean sea-ice extent for decadal periods. Note that this is the opposite conclusion reached above regarding sea-ice extent trends.

In the cascade (Figure 4), we calculated a measure of inter-model spread. Over 2046 to 2065, the inter-model range was 3.5 million km^2 across the six GCMs considered (9.4 million km^2 across the full CMIP5 ensemble). For each model, the mean extent was computed by averaging over the 20 year period and over all available realizations, and the quoted inter-model range was computed from these model mean extents for each RCP scenario. The inter-model ranges were 4.2, 3.8 and 2.5 million km^2 in RCP2.6, RCP4.5 and RCP8.5 respectively . The inter-model ranges across the full CMIP5 ensemble were 9.9, 9.4 and 8.8 million km^2 in RCP2.6, RCP4.5 and RCP8.5 respectively.

Since the ensemble sizes of the models used in the cascade were small (3 to 6 realizations), inter-realization spread could influence the calculation of the mean extents for each model and hence the inter-model range. Using CESM1-LE (RCP8.5) we calculated the effect of ensemble size on 20 year mean sea-ice extent for different periods, and found it to be small (Figure S7). Indeed, inter-realization range in mean extent between the realizations of a single model over 2046 to 2065, are 0.7, 0.6 and 0.5 million km^2 in RCP2.6, RCP4.5 and RCP8.5 respectively when averaged over all CMIP5 models. In individual models the inter-realization spread reached up to maximum values of 1.5, 1.5 and 2.1 million km^2 in RCP2.6, RCP4.5 and RCP8.5 respectively across the full CMIP5 ensemble. The inter-realization spreads are thus much smaller than the inter-model spreads of around 9 to 10 million km^2 quoted above. The scenario uncertainty was determined by computing the multi-model mean extent for each of the three RCP scenarios, and then calculating the range, which is 1.3 million km^2 across the CMIP5 ensemble over 2046 to 2065. In conclusion, over 2046 to 2065, inter-model spread is much larger than inter-realization spread and inter-scenario spread, and being restricted to a small sample size in the cascade does not influence that conclusion.

References

[1] Fetterer, F., and Knowles, K. and Meier, W. and Savoie, M.. Sea Ice Index, updated daily. Boulder, Colorado USA: National Snow and Ice Data Center. (2002). http://dx.doi.org/10.7265/N5QJ7F7W.

[2] Huber, M. & Knutti, R. *Nature Geosci* **7**, 651–656 (2014).

[3] Fyfe, J. C., Gillet, N. P. & Zwiers, F. W. *Nature Clim. Change* **3** (2013).

DOI: 10.1038/NCLIMATE2483

Table S1: List of models used in this study, and the number of realizations per model for the RCP2.6, 4.5 and 8.5 experiments (in each case appended to the historical experiment). Models with multiple realizations for each RCP which were used in the cascade of Figure 3 are indicated with a *.

Model	**RCP2.6**	**RCP4.5**	**RCP8.5**
ACCESS1-0	0	1	1
ACCESS1-3	0	1	1
BCC-CSM1-1	1	1	1
BNU-ESM	1	1	1
CCSM4*	6	6	6
CESM1-BGC	0	1	1
CESM1-CAM5*	3	3	3
CMCC-CM	0	1	1
CMCC-CMS	0	1	1
CNRM-CM5	1	1	5
CSIRO-Mk3-6-0	10	10	10
CanESM2	5	5	5
EC-EARTH	1	10	10
FGOALS-g2	1	1	1
GFDL-CM3	1	2	1
GFDL-ESM2M	1	1	1
GISS-E2-H	3	15	3
GISS-E2-R	3	17	3
HadGEM2-CC	0	1	0
HadGEM2-ES*	3	3	3
IPSL-CM5A-LR*	4	4	4
IPSL-CM5A-MR	1	1	1
IPSL-CM5B-LR	0	1	1
MIROC-ESM	1	1	1
MIROC-ESM-CHEM	1	1	1
MIROC5*	3	3	3
MPI-ESM-LR*	3	3	3
MPI-ESM-MR	1	3	1
MRI-CGCM3	1	1	1
NorESM1-M	1	1	1
NorESM1-ME	1	1	1
Total	57	102	76

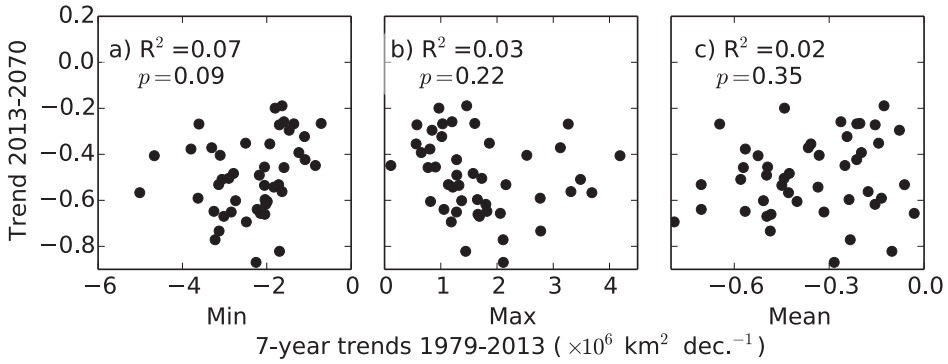

Figure S1: **Historical and future trends.** The relationship between the a) minimum, b) maximum and c) mean 7 year trend over 1979 to 2013 and the future trend over 2013 to 2070 in the CMIP5 ensemble historical-RCP4.5 experiment. The coefficient of determination, R^2, of the linear regression between each set of variables are given, along with the p values for the test of the null hypothesis that the slope of the regression line is equal to zero.

DOI: 10.1038/NCLIMATE2483

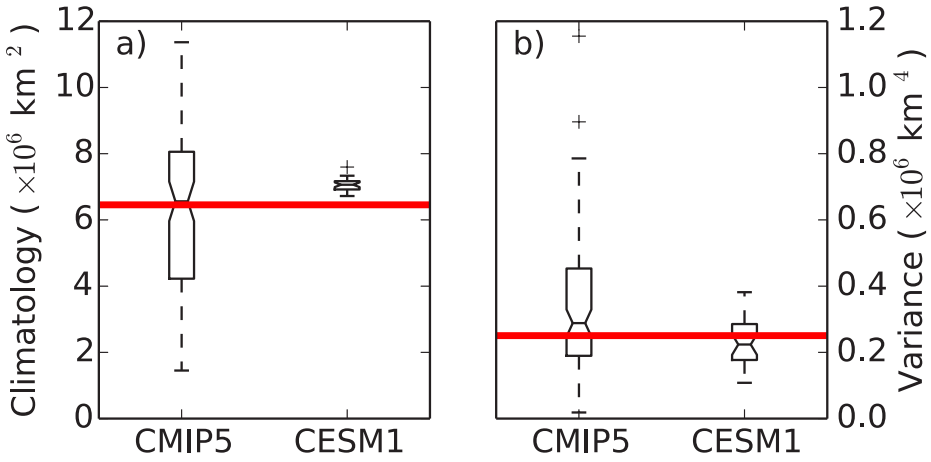

Figure S2: **Arctic sea-ice climatology and variance.** Notched box plots of a) mean September sea-ice extent (climatology) and b) linearly detrended variance in September sea-ice extent over 1980 to 2010 for the 102 realizations from CMIP5 and 30 realizations from CESM1 LE. The solid red lines indicates the observed values over this period.

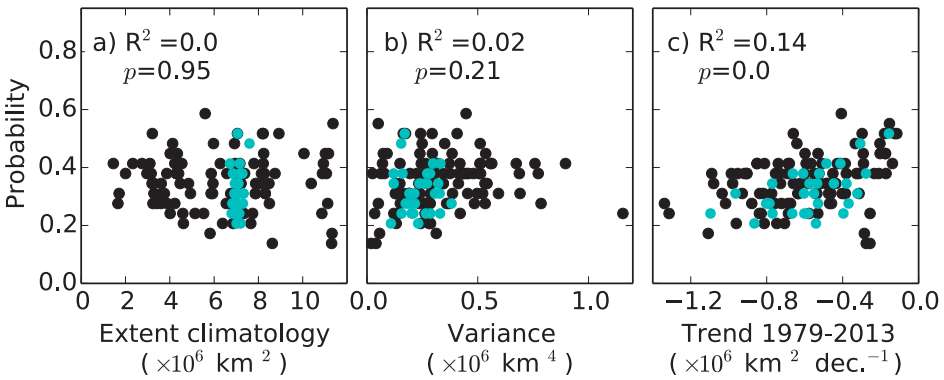

Figure S3: **Relationships between pause likelihood and mean state.** The relationship between a) historical September Arctic sea-ice climatology, b) sea-ice variance and c) long term trend over 1979 to 2013 with the probability of a 7 year pause in sea-ice extent loss. The coefficient of determination, R^2, of the linear regression between each set of variables are given, along with the p values for the test of the null hypothesis that the slope of the regression line is equal to zero. Black dots are for CMIP5, cyan dots are for CESM1 LE.

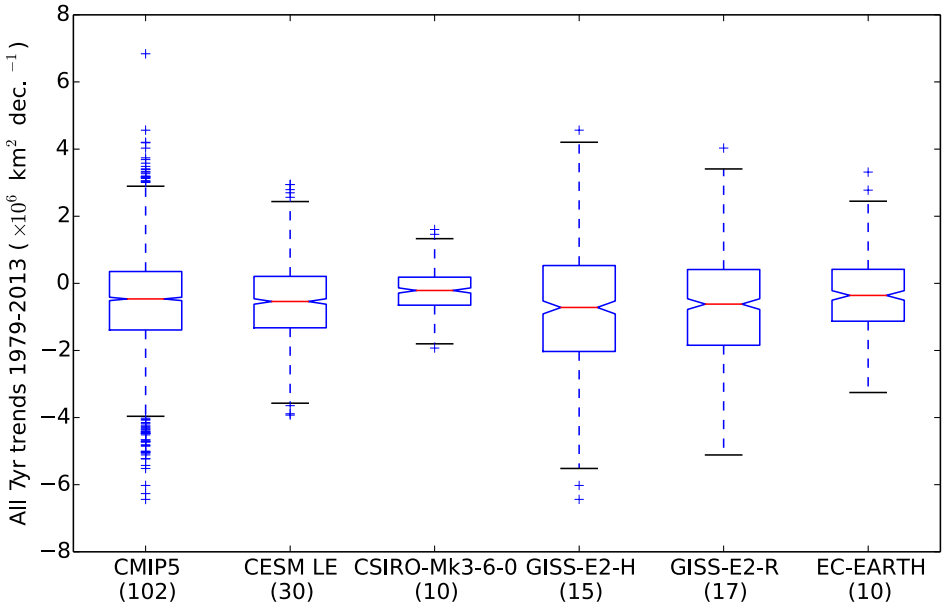

Figure S4: **Spread of 7 year trends across various models.** The spread is shown in notched box plots of all 7 year trends between 1979 and 2013 for the CMIP5 ensemble and for five individual models. Whiskers extend from the 25^{th} to 75^{th} percentile. The number of realizations is indicated below the model name in parentheses. It should be noted that the two GISS models comprise realizations with perturbed physics and realizations with perturbed initial conditions, while the other three models comprise only members with perturbed initial conditions.

DOI: 10.1038/NCLIMATE2483

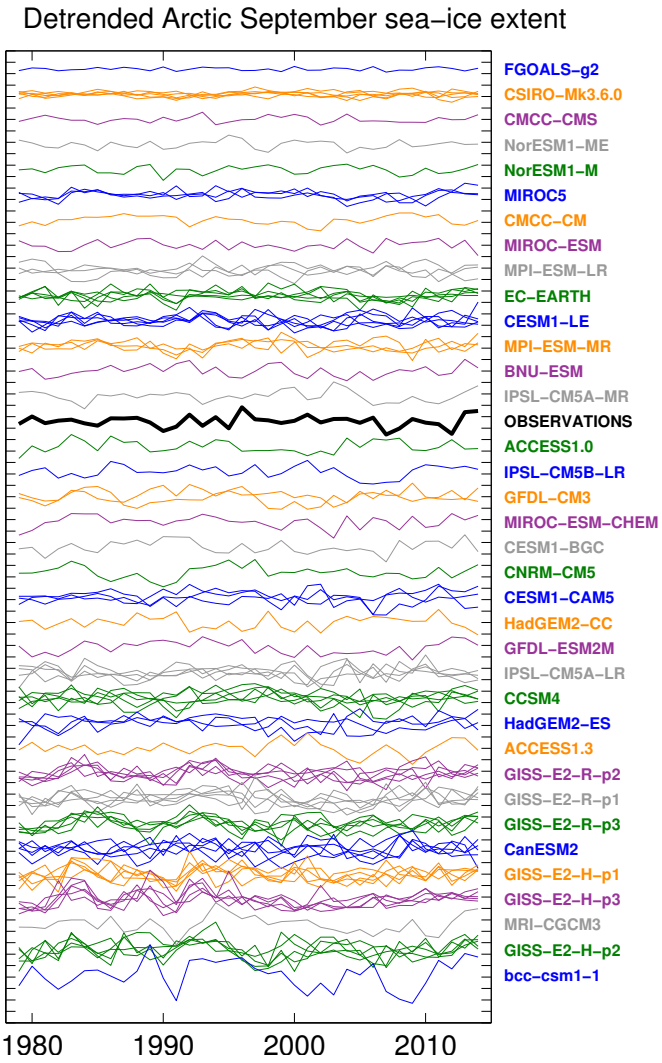

Figure S5: **Sea-ice extent variability in CMIP5.** The detrended Arctic September sea-ice extent in the CMIP5 and CESM1 LE models, as well as observations, ordered in terms of their variance from lowest (top) to highest (bottom). A quadratic detrending was applied, and for models up to the first five ensemble members are used, with ordering done based on variance averaged over each models runs. Note that the ordering may change given a larger number of realizations for each model, since some spread in the variances is inter-realization (not inter-model) based. One example of that is shown by the CESM1-CAM5 and CESM1-LE, which are from the same model, but show different variances if only the first five realizations are used to order the models.

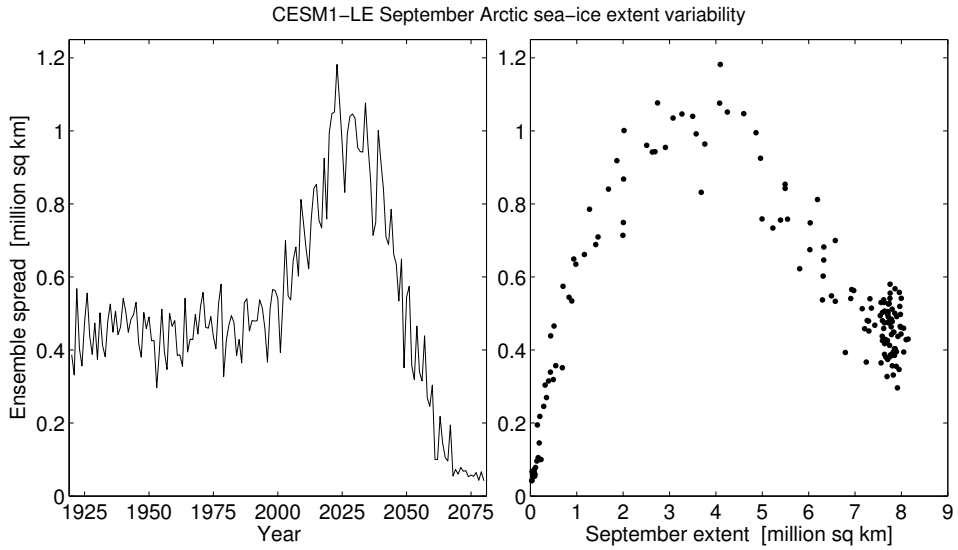

Figure S6: **Variability changes in CESM1 LE.** The ensemble spread (standard deviation) of September sea-ice extent in the CESM1 Large Ensemble as a function of time (left) and as a function of ensemble mean sea-ice extent (right).

DOI: 10.1038/NCLIMATE2483

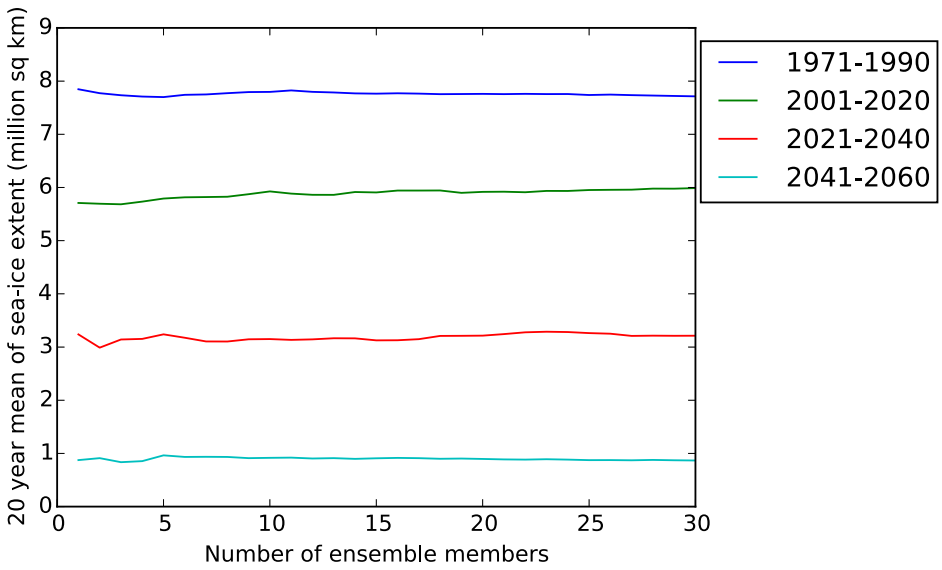

Figure S7: **Mean extent as a function of ensemble size.** The mean sea-ice extent over various 20 year periods in the CESM1 Large Ensemble as a function of the number of ensemble members used.