Arctic Inversion Strength in Climate Models

BRIAN MEDEIROS, CLARA DESER, ROBERT A. TOMAS, AND JENNIFER E. KAY

National Center for Atmospheric Research,* Boulder, Colorado

(Manuscript received 2 August 2010, in final form 14 January 2011)

ABSTRACT

Recent work indicates that climate models have a positive bias in the strength of the wintertime low-level temperature inversion over the high-latitude Northern Hemisphere. It has been argued this bias leads to underestimates of the Arctic’s surface temperature response to anthropogenic forcing. Here the bias in inversion strength is revisited. The spatial distribution of low-level stability is found to be bimodal in climate models and observational reanalysis products, with low-level inversions represented by a stable primary mode over the interior Arctic Ocean and adjacent continents, and a secondary unstable mode over the Atlantic Ocean. Averaging over these differing conditions is detrimental to understanding the origins of the inversion strength bias. While nearly all of the 21 models examined overestimate the area-average inversion strength, conditionally sampling the two modes shows about half the models are biased because of the relative partitioning of the modes and half because of biases within the stable mode.

1. Introduction

Low-level temperature inversions are a noted feature of the Arctic winter climate (Serreze et al. 1992; Zhang et al. 2011). The so-called Arctic inversion mediates the surface energy balance and contributes to amplifying the high-latitude surface temperature response to anthropogenic increases in greenhouse gas (GHG) concentrations (Serreze and Barry 2005). The amplified warming over the Arctic Ocean and surrounding continents in recent years has been most pronounced in the lower troposphere and stably stratified PBL (Serreze et al. 2009; Screen and Simmonds 2010). As Arctic sea ice and high-latitude terrestrial snow cover diminish in response to increasing GHG, the inversion is expected to weaken, with consequences for the rate of surface warming, cloud type and amount, and other effects (Pavelsky et al. 2010; Deser et al. 2010; Alexander et al. 2010; Kay and Gettelman 2009). Boé et al. (2009) point out that climate models tend to overestimate the Arctic inversion strength, some by more than a factor of 2, and suggest that this bias leads to an excessive negative longwave radiative feedback and hence reduced climate sensitivity. Thus, the wintertime low-level temperature inversion is of key interest for understanding high-latitude climate change.

The Arctic inversion that caps the PBL is influenced by large-scale advection and local processes, particularly surface fluxes and entrainment (Busch et al. 1982). The inversion exhibits pronounced seasonality: elevated inversions are common during spring and summer when low clouds are often present, while surface-based inversions are pervasive during autumn and winter (Tjernström and Graversen 2009). Boé et al. (2009) and Pavelsky et al. (2010) report winter inversion strengths <3 K for the Arctic as a whole (northward of 70° and 64°N, respectively), based on several reanalysis and satellite products. This value is considerably smaller than the ≈10-K inversion strength obtained by Tjernström and Graversen (2009) from radiosonde observations over the central Arctic Ocean from the Surface Heat Budget of the Arctic Ocean (SHEBA) experiment.

Here we analyze the spatial distribution of the mean wintertime inversion strength over high latitudes of the Northern Hemisphere. We compare twentieth-century integrations from 21 climate models (Table 1) in the Third Coupled Model Intercomparison Project (CMIP3) archive (Meehl et al. 2007) with observational estimates from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; used in Boé

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

**Corresponding author address:** Brian Medeiros, NCAR/CGD, P.O. Box 3000, Boulder, CO 80307-3000. E-mail: brianpm@ucar.edu

DOI: 10.1175/2011JCLI3968.1

© 2011 American Meteorological Society
et al. 2009; Pavelsky et al. 2010) and the updated ERA-Interim Reanalysis (Simmons et al. 2007). For ERA-40, we show results using the period with the highest-quality remote sensing products in the assimilation (1987–2001; Uppala et al. 2005), though using the full period (1957–2001; not shown) supports all conclusions presented; for ERA-Interim, we use the full record (1989–2009). For the reanalysis and CMIP3 models, monthly means are analyzed. Additional simulations with the National Center for Atmospheric Research (NCAR) Community Atmosphere Model, version 3 (CAM3), are used to investigate the temporal characteristics of the inversion.

2. Arctic inversion strength

The Arctic inversion strength can be estimated by differencing temperatures at a pressure level above the inversion and near the surface: \( \Delta T = T_{atm} - T_{sfc} \). The 850-hPa level is chosen as the upper level, which is expected to be above the PBL (Serreze et al. 1992). The lower temperature is often taken at 1000 hPa, but wintertime surface pressure deviates enough from 1000 hPa that surface temperature is preferred here; the main consequence of this choice is to slightly increase \( \Delta T \) over the ocean. Whatever definition is employed, this is a crude estimate of the inversion strength, but is appropriate for comparing the stability of the lower troposphere in climate models, which have coarse vertical resolution. A strict definition would exclude \( \Delta T \), since such unstable conditions do not represent an inversion, but the neglect of this criterion by many recent papers prompts us to examine the consequences of including unstable conditions in the analysis.

Figure 1 shows the average \( \Delta T \) for the 21 years of ERA-Interim reanalysis and the ensemble of CMIP3 models (all years). Because the Arctic is data sparse, the reanalysis largely reflects the underlying model, but the inversion in the ERA-40 reanalysis compares reasonably well to radiosonde observations (Tjernström and Graversen 2009; Zhang et al. 2011). Both reanalysis products are
interpolated to a 2.5° latitude–longitude grid, a resolution similar to the models, and only the winter season (November–February) is sampled. The ERA-Interim map shows that the Arctic Ocean has a monthly-mean $\Delta T$ of $\sim$6 K, higher near North America and lower near Asia; the multimodel mean shows a similar pattern, but stronger $\Delta T$. The North Atlantic, however, shows unstable conditions, exceeding $\sim$9 K off the Norwegian coast, and the reanalysis has more unstable values than the models. Strong inversions are also evident over land, especially Siberia, and the reanalysis and models agree to a large extent, except near Scandinavia. The unstable conditions over the ocean are associated with little to no sea ice, as can be seen from the corresponding sea ice concentration distributions in the bottom panels of Fig. 1.

The black curves in Fig. 2 show histograms of $\Delta T$ in both reanalysis products for ocean and land using the same domain as in Fig. 1 [64°–90°N, which we loosely define as the “Arctic” to be consistent with recent work; e.g., Pavelsky et al. (2010)]. The histograms are weighted...
by area and comprise all monthly-mean values at each grid point in the domain. The oceanic distribution is bimodal, with maxima around 6 and 29 K. Figure 1 suggests that the primary mode represents the Arctic Ocean while the secondary mode corresponds to the North Atlantic. The terrestrial histogram (excluding Greenland) is unimodal, centered on stable values of $D_T$ around 7 K. There is a tail on the land distribution toward unstable conditions, and these tend to be focused on Scandinavia, perhaps showing the influence of North Atlantic conditions.

The CMIP3 models are also included in Fig. 2, based on data availability (Table 1). We exclude the Institute of Atmospheric Physics (IAP) Flexible Global Ocean–Atmosphere–Land System Model (FGOALS) from most figures because its excessive sea ice makes it an outlier (cf. Zhang and Walsh 2006). The models show the same bimodal pattern as the reanalysis products. The unstable, North Atlantic mode tends to be less common, but peaks at a more unstable value. In the Arctic, the models show a variety of solutions, but tend to overestimate the frequency of stable conditions compared with the reanalysis products. Eleven models have stable modes more stable than the reanalysis, 5 are within 1 K of the reanalysis mode, and 5 have a less stable mode. Thus this analysis confirms the finding that climate models overestimate the Arctic inversion strength, but also demonstrates that the bias stems from different distributions among the models. The land distributions show more agreement between models and reanalysis, but most models still exaggerate the stability and underestimate the unstable portion of the reanalysis distributions.

Figure 3 shows the average oceanic value of $D_T$ for the region poleward of 64°N (gray). Dividing the distributions at $D_T = 0$ produces two well-separated “regimes,” showing stable (blue) and unstable (red) conditions. The word “regime” is used here to distinguish two sets of behavior associated with differing surface conditions and PBL structure, but not necessarily unique physical mechanisms. All the models (except IAP FGOALS) exhibit both stable and unstable regimes, with the stable one being the larger contribution to the regional average. The ERA-Interim has slightly smaller regionally averaged $D_T$ than the ERA-40, and also within the stable part of the distribution. Most models place the Arctic average value within 1σ of either reanalysis product.

Figure 3 shows that the apparent agreement between the models and reanalysis arises from compensating biases in the partitioning of the regimes and the magnitude of $D_T$ within the regimes. Within the stable regime, about half of the models have average $D_T$ more than 1σ larger than the reanalysis values. Agreement is better in the unstable regime, where almost all the models are within the reanalysis range. This comparison shows that the partitioning between regimes significantly impacts the average $D_T$. Take for example the Institute of Numerical Mathematics Coupled Model, version 3 (INM CM3): the regional average $D_T$ is close to the reanalysis value, but its stable regime is more stable; the large area covered by the unstable regime in this model reduces the regional average $D_T$ to a value close to the reanalysis value. This is corroborated by the model’s placement to the left side of Fig. 3, indicating it has relatively little sea ice.

Because the distribution of $D_T$ is bimodal, the mean value can be strongly biased and has limited utility. The presence of extremes in either mode can further bias the mean. To test for such bias, we have repeated the analysis above using the area-weighted median (not shown). Doing so removes the impact of extreme values, but retains the effects of different areal coverage of the regimes. The Arctic area average (gray in Fig. 3) is consistently smaller than the median because the stable regime covers a larger...
area than the unstable regime in the winter. Within the individual stable and unstable regimes, however, the median and mean agree well.

The regional-average $\Delta T$ is partly determined by the relative sizes of the regimes, and the partitioning varies across models. Therefore, care is required for interpreting regional averages. Figure 4 shows how the choice of domain affects the regional average. Including substantial subpolar ocean area results in negative average $\Delta T$. For ERA-40, it is only north of 60°N that average $\Delta T$ is positive and the ice-covered Arctic overwhelms other signals; for ERA-Interim the sign change occurs around 62°N, while the CMIP ensemble average becomes positive at about 55°N. Both ERA-40 and ERA-Interim show a stable $\Delta T$ over land that increases as the domain contracts northward until the land area becomes small (~70°N).

The excessively strong Arctic inversion appears in most of the models even for large domains (Fig. 4). Only two models are consistently less stable than the reanalysis over the ocean. The models are more uniformly distributed about the reanalysis over land, but the ensemble mean is more stable than either reanalysis. The variation in the model spread with latitude reinforces that some conclusions could depend on one’s definition of the Arctic region. Conditioning the analysis on regimes—using $\Delta T$ as is done here or by using other criteria—removes such sensitivity; the averaging of Fig. 4 for the stable regime demonstrates negligible latitude dependence because lower latitudes contribute little to the stable regime (not shown). Conditional sampling thus provides more robust results than averaging over multiple regimes in a specified geographical area.

The importance of spatial sampling raises the question of whether temporal sampling also matters. To explore some issues of temporal sampling, we analyze a CAM3 integration forced by the observed evolution of SST and sea ice concentration and thickness during 1980–2008. Figure 5 shows the distributions of $\Delta T$ constructed from 4-times-daily instantaneous fields (black), 5-day averages of the instantaneous output (orange), and monthly mean values (from all time steps, blue). For the oceanic distributions, the 5-day average and the monthly mean distributions are both bimodal as in Fig. 2, with the 5-day averages having slightly larger variability. The distributions based on instantaneous data, however, contain less evidence for bimodality. The interval between the stable and unstable modes near $\Delta T = 0$ appears to be
populated in the instantaneous fields mostly at the expense of the stable regime, and there is also enhanced frequency of very stable conditions. This additional variability represents transient weather conditions, which appear to have a time scale of less than 5 days. Over land areas, the instantaneous fields also show a broader distribution, and contrasting with the oceanic distributions, actually appear more bimodal than the lower-frequency sampling.

Figure 5 also shows the distributions of monthly means from the Community Climate System Model, version 3 (CCSM3), and ERA-Interim from Fig. 2. The atmosphere-only and fully coupled results are similar to each other, particularly over land (SST and sea ice conditions differ between the integrations, precluding direct comparison). In the stable mode, CAM3 is shifted toward slightly more stable oceanic conditions than CCSM3, and both are more stable than the reanalysis. The CCSM3 shows a sharp peak near −9 K compared with the broader peak in CAM3 centered at about −15 K, but the area covered by the unstable regime is comparable. The overall similarity between CAM3 and CCSM3 suggests that for this model atmospheric processes are probably responsible for the bias in $\Delta T$ compared with reanalysis. This view is also supported by Fig. 3, which shows that the CCSM3 regional average is strongly influenced by its large $\Delta T$ in the stable regime.

3. Summary and discussion

Most of the CMIP3 climate models overestimate the stability of the lower troposphere over the Northern Hemisphere high latitudes. This agrees with previous studies and elaborates on them by exploring the spatial distribution of $\Delta T$. Differences in surface conditions and large-scale circulation lead to a bimodal distribution of
ΔT over the ocean. The stable mode represents stable boundary layers over cold surfaces (e.g., sea ice and snow), while the unstable mode is found over open water, mostly in the North Atlantic, and represents more well-mixed boundary layers. When these modes are treated as separate regimes the disagreement with reanalysis remains, though many models are within the variability of the reanalysis. In the stable regime, about half the models overestimate the mean inversion strength by more than the reanalysis 1σ level.

This analysis makes clear that the reported bias in Arctic inversion strength in CMIP3 models arises from two different sources of error. First, the relative area covered by stable versus unstable conditions. Second, the representation of the lower troposphere (i.e., the PBL) within those conditions, especially the representation of stable boundary layers. These are proximate factors, but determining the ultimate causes of the bias for each model requires additional analysis. Investigating the sea ice distribution and large-scale atmospheric circulation patterns may help explain the partitioning of stable and unstable conditions. Diagnosing problems in the representation of stable boundary layers is likely impossible based on the monthly-mean CMIP3 archive. This is mainly because monthly time scales conflate many separate processes affecting the PBL. For example, advection of warm air over cold surfaces is often associated with transient conditions, and separating such transport factors from stable conditions driven by radiative cooling is not possible in monthly averages. A secondary issue is that stable boundary layers are often very shallow, with complicated vertical structure (Mahrt 1999), and climate models do not properly resolve such details. As such, detailed understanding of the stable boundary layer will require intricate knowledge of a model’s parameterizations and implementation, and so be model dependent. It is possible that future comparisons (e.g., CMIP5) that compile high-frequency output will allow more nuanced understanding of the origin of the Arctic inversion strength biases.

Differences between the ERA-Interim and ERA-40 products are relatively small, justifying use of either one for comparison with the models. The ERA-40 reanalysis has previously been favorably compared with observations (e.g., Tjernström and Graversen 2009; Zhang et al. 2011), and previous analyses suggest the present conclusions would stand with other reanalysis products (Boé et al. 2009). Several studies use Atmospheric Infrared Sounder (AIRS) temperature profiles to investigate the Arctic inversion (e.g., Kay and Gettelman 2009; Devasthale et al. 2010); our analysis of AIRS inversion strength (not shown) yields results similar to the comparison with reanalysis (see also Pavelsky et al. 2010). We find, however, that AIRS surface (and 1000 hPa) temperature is biased warm compared with reanalysis; this weakens ΔT and precludes quantitative comparison to the CMIP3 models. We have investigated the impact of using monthly averages by using 4-times-daily instantaneous output from the CAM3. The high-frequency output contains weather noise at short time scales, but temporal averages quickly converge toward the monthly mean distribution. We have explored the sensitivity of the CMIP3 model results and the ERA-40 reanalysis to the time period chosen for analysis and found little impact on our conclusions with regard to the time-mean inversion strength.

These results provide some context for the exaggerated Arctic inversion strength in climate models by demonstrating that some models overestimate local inversion strength and others contain spatial biases. The central Arctic in winter is completely ice covered, while the North Atlantic is largely ice free. Over ice, stable boundary layers and strong inversions are expected, as ice buffers ocean–atmosphere heat exchange. In the North Atlantic, surface fluxes can be large as warm water is advected northward to meet cold, polar air. This drives a deeper, more well-mixed PBL. The mean Arctic inversion strength is determined by the partitioning between these PBL structures, which changes as sea ice decreases (Kay and Gettelman 2009; Deser et al. 2010). Disagreement among climate models is also affected by the representation of physical processes, such as turbulent mixing, surface fluxes, and clouds. Isolating regimes, using ΔT or other criteria, provides for more robust model evaluation and should help elucidate the origin of biases. Here we find that about half the models overestimate inversion strength even within stable conditions, suggesting deficiencies in representing stable boundary layers in those models. Some of the other models may better represent the inversion, but contain spatial biases (possibly related to sea ice or large-scale atmospheric circulation patterns) that impact the regional inversion strength. Focusing on the PBL provides a link between the surface energy budget and inversion strength and represents an approach to evaluating the representation of high-latitude boundary layer processes in models.

Acknowledgments. We thank the anonymous reviewers for their constructive comments. We thank D. Seidel, Y. Zhang, C. Ao, J. C. Golaz, and S. Park for encouraging this work, and PCDI and the Working Group on Coupled Modelling for making available the WCRP CMIP3 dataset with support from the Office of Science, U.S. Department of Energy. C. Deser and R. Tomas acknowledge support from the NSF Office of Polar Programs. B. Medeiros was supported by the Office of Science.
REFERENCES


