

Decadal predictability of tropical basin average and global average sea surface temperatures in CMIP5 experiments with the HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 global Earth System Models

Vikram M. Mehta,¹ Hui Wang,^{1,2} and Katherin Mendoza¹

Received 14 December 2012; revised 5 February 2013; accepted 6 February 2013; published 14 June 2013.

[1] This paper describes initial results from a broadscale study to assess decadal climate hindcast skills of the HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 global Earth System Models (ESMs) in experiments conducted under the Coupled Model Intercomparison Project 5. Analyses of decadal hindcast and simulation experiments using historical aerosol optical depths show statistically significant decadal predictability skill of global average and tropical sea surface temperature (SST) anomalies during 1961–2010. The skill, however, varies by averaging region and decade. It was also found that volcanic eruptions influence SSTs and are one of the sources of decadal SST hindcast skill. In the actual climate system, however, volcanic eruptions themselves are not predictable, and therefore, their effects on the climate system can only be predicted after eruptions. In the four ESMs utilized in this study, decadal hindcast skills of SST anomalies over ocean basin size averaging regions generally improve due to model initialization with observed data. **Citation:** Mehta, V. M., H. Wang, and K. Mendoza (2013), Decadal predictability of tropical basin average and global average sea surface temperatures in CMIP5 experiments with the HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 global Earth System Models, *Geophys. Res. Lett.*, 40, 2807–2812, doi:10.1002/grl.50236.

1. Introduction

[2] Skillful decadal climate predictions can greatly benefit planning in many societal sectors, such as agriculture, reservoir operations, forest fires, municipal water supply and drainage systems, fisheries and wildlife hatcheries, hydroelectricity generation, thermal and nuclear power plant operations, river- and reservoir-based recreation industry, transportation, and state and national government decisions [Mehta *et al.*, 2013]. Prospects for decadal climate prediction depend on prospects for obtaining skillful predictions/projections of interannual variability such as El Niño–Southern Oscillation (ENSO), natural decadal climate variability (DCV), climate system responses to variations in solar particulate and radiative emissions and volcanic eruptions, and responses to human-induced changes in land use-cover and atmospheric constituents. There

have been three decadal hindcast (retrospective forecast) studies with Earth System Models (ESMs) [Smith *et al.*, 2007; Keenlyside *et al.*, 2008; Pohlmann *et al.*, 2009]. In these studies, ESMs were initialized from observed data, as in seasonal climate forecasting, and natural and anthropogenic changes in aerosol optical depth (AOD) were prescribed from observations-based estimates (or scenarios) as in climate change experiments. These initial hindcast studies demonstrated enhanced skill from initialization on a global scale [Smith *et al.*, 2007] and over the North Atlantic [Keenlyside *et al.*, 2008; Pohlmann *et al.*, 2009].

[3] Following these initial decadal climate predictability studies, the World Climate Research Program organized the Coupled Model Intercomparison Project 5 (CMIP5) to assess the ability of the current generation of ESMs used in climate and impacts assessments by the Intergovernmental Panel on Climate Change to simulate and hindcast decadal climate. Two sets of core decadal prediction experiments are being conducted under CMIP5 [Meehl *et al.*, 2009]. The first set is a series of 10 year hindcasts starting approximately in 1970, 1990, and 2000. The second is a series of 30 year hindcasts starting in 1960, 1980, and 2005, the last a combined hindcast-forecast. In both sets, AODs (including those due to volcanic eruptions) and solar radiation are prescribed from past observations. Each experiment has a minimum ensemble size of three members. These experiments are somewhat idealistic and exploratory, especially in view of the well-known difficulty of predicting volcanic eruptions well in advance.

[4] The results reported here are a part of a broadscale study to assess simulation and decadal hindcast skills of some of the ESMs participating in CMIP5. The following questions are addressed in this paper: Is there any significant decadal hindcast skill of global and regional average sea surface temperature (SST) anomalies? If yes, what are the relative roles of initial (model integrations starting from observed data at a specific time) and boundary (prescribed atmospheric constituents and solar radiation) conditions in hindcast skill? Do changes in AOD due to volcanic eruptions influence decadal SST predictability? If yes, can the role of this influence be quantified in terms of changes in prediction skill? Skills of these ESMs to simulate and hindcast DCV phenomena such as the Pacific Decadal Oscillation, the tropical Atlantic SST gradient variability, the ENSO variability, and tropical warm pools' variability will be described elsewhere.

2. CMIP5 and Observed Data Sets

[5] We used SST and AOD data from the U.K. Meteorological Office Hadley Centre HadCM3 model, the U.S.

¹Center for Research on the Changing Earth System, Catonsville, Maryland, USA.

²Wyle ST&E Group and NOAA/Climate Prediction Center, Camp Springs, Maryland, USA.

Corresponding author: Vikram M. Mehta, Center for Research on the Changing Earth System, 5523 Research Park Drive, Suite 315, Catonsville, MD 21228, USA. (vikram@crcees.org)

Geophysical Fluid Dynamics Laboratory (GFDL) CM2.1 model, the U.S. National Center for Atmospheric Research (NCAR) CCSM4 model, and the Model for Interdisciplinary Research on Climate 5 (MIROC5) from Japan. Table 1 summarizes some major attributes of these models and the CMIP5 experiments carried out with them. In the CMIP5 hindcast experiments, the GFDL CM2.1 used a fully coupled initialization scheme [Zhang *et al.*, 2007], the MIROC5 used an ocean-only initialization scheme [Tatebe *et al.*, 2012], the NCAR-CCSM4 used ocean and sea ice initial conditions from a historical forced experiment [Yeager *et al.*, 2012], and the HadCM3 was initialized by relaxation to analyzed ocean and atmosphere observations [Smith *et al.*, 2007]. In all CMIP5 experiments, Northern Hemisphere and Southern Hemisphere time series of AOD, based on observations (Ammann *et al.* [2003] in the NCAR ESM and Sato *et al.* [1993] and Hansen *et al.* [2002] in the other three ESMs), were specified. These data sets provide zonal average, vertically resolved AOD for visible wavelengths and column average effective radii of aerosols [Stenchikov *et al.*, 2006].

[6] We used the Extended Reconstructed SSTs (ERSST) [Reynolds *et al.*, 2002] from 1961 to 2010 for comparison with simulated and hindcast SSTs.

3. Analysis Techniques

[7] Following Smith *et al.* [2007], Keenlyside *et al.* [2008], and Pohlmann *et al.* [2009], we estimated decadal hindcast skill in the form of root mean square hindcast errors and correlation coefficients between hindcast and observed variables. We also estimated simulation skill in the form of root mean square simulation errors and correlation coefficients between simulated and observed variables. The skill estimates were evaluated based on the monthly average ensemble data from each model and also the data from a multimodel ensemble (MME) [Krishnamurti *et al.*, 2000]. The latter was simply an average of the average ensemble data from each model. In this way, each model was treated equally in the MME. Prior to calculating correlation coefficients, all data were detrended over the 1961–2010 period. The Monte Carlo technique [e.g., Wilks, 1995] was used to estimate statistical significance of correlation coefficients. Only hindcast skill as indicated by correlation coefficient is described in this paper. Correlation coefficients equal to or greater than 95% confidence limit are referred to as

statistically significant in this paper. Also, negative correlation coefficients are referred to as negative skill.

4. Decadal Hindcast Skill of Global and Regional SSTs

[8] Quantitative assessments of decadal hindcast skill of regionally averaged SSTs by the four ESMs and by the MME are described in this section. SST anomaly correlations at each grid point of hindcasts by each ESM over 10 year hindcast periods strongly indicated (not shown) that hindcast skills are generally homogeneous within each tropical ocean basin, so SST anomalies were averaged within each tropical ocean basin and hindcast skills of regionally averaged SST anomalies were further studied. Correlation coefficients were estimated between hindcast and observed monthly SST anomalies over each hindcast (10 years) period; the anomalies were averaged globally and in the near-tropical (20°S–20°N) latitude belt, the tropical Indian Ocean, the tropical Pacific Ocean, and the tropical Atlantic Ocean. These correlation coefficients, with 95% and 99% confidence limits, are shown in Figure 1. The hindcast skill for regionally averaged SST anomalies over the entire 1961–2010 period, shown in Figure 1a, is significant for global average anomalies for all models, followed by the tropical Indian Ocean and the tropical Atlantic Ocean. Figure 1a also shows that MIROC5 has significant skill in all averaged regions and that HadCM3 and CCSM4 have an insignificantly small negative skill in the tropical Pacific over the 50 year period.

[9] The decade-by-decade hindcast skills of regionally averaged SST anomalies are shown in Figures 1b to 1f. In general, hindcast skill fluctuates from decade to decade in all regions. The hindcast skill of globally averaged SST anomalies (Figure 1b) fluctuates from decade to decade, but MIROC5 has significant skill in all decades. It is somewhat counterintuitive that the highest skill is in the 1960s when the observed data, especially ocean data, were sparsest of the entire 50 year period. This is addressed further in section 5.

[10] The basin average skill in the tropical ocean belt (Figure 1c) is significant only in MIROC5 in 1960s, 1970s, 1980s, and the first decade of the 21st century (“the oughts”); in CCSM4, GFDL-CM2.1, and HadCM3 in the 1970s; and in GFDL CM2.1 in the oughts. A peculiar observation from Figure 1c is that the skill between observed and HadCM3-hindcast SST anomalies is significantly negative

Table 1. CMIP5 Models Used in This Study

Model	Institute	Experiment	Ensemble Members	SST Resolution
GFDL-CM2.1	NOAA Geophysical Fluid Dynamics Laboratory, U.S.A.	Decadal hindcast (1960, 1970, 1980, 1990, 2000), historical runs (1861–2040)	10 (decadal) 10 (historical)	1° (long.) × 0.34° (lat.) at Eq. and 1° (lat.) at 28° and poleward
HadCM3	Hadley Centre, U.K.	Decadal hindcast (1960, 1970, 1980, 1990, 2000), historical runs (1860–2005)	10 (decadal) 10 (historical)	1.25° × 1.25°
MIROC5	Atmosphere and Ocean Research Institute (Univ. of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	Decadal hindcast (1960, 1970, 1980, 1990, 2000), historical runs (1859–2012)	6 (decadal) 4 (historical)	Rotated pole grid ~ 1.41° (long.) × 0.79° (lat.)
CCSM4	National Center for Atmospheric Research, U.S.A.	Decadal hindcast (1960, 1970, 1980, 1990, 2000), historical runs (1961–2005)	10 (decadal) 6 (historical)	1.25° × 1.25°

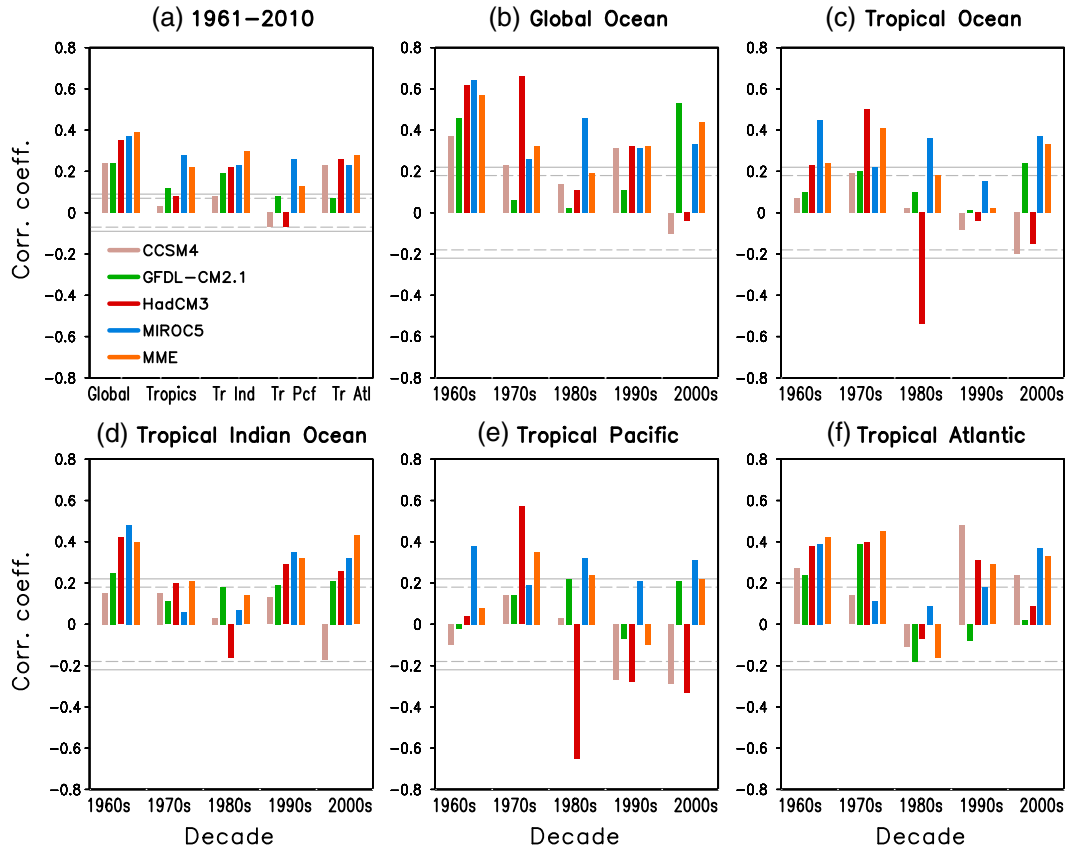


Figure 1. Correlations between ERSST and CMIP5 decadal hindcast SST anomalies averaged regionally for (a) the 1961–2010 period and for each decade over (b) global, (c) tropics, (d) tropical Indian Ocean, (e) tropical Pacific Ocean, and (f) tropical Atlantic Ocean. Solid (dash) gray lines denote the 99% (95%) significance level.

in 1980s and negative but insignificantly small in 1990s and the oughts; CCSM4 also has negative skill in the oughts. The MME shows significant skill in 1960s, 1970s, 1980s, and the oughts. It is remarkable that there is no significant skill in any of the ESMs or the MME in the tropical ocean average in 1990s. GFDL CM2.1, MIROC5, and HadCM3 have significant hindcast skill in 1960s, 1990s, and the oughts of average SST anomalies in the tropical Indian Ocean (Figure 1d).

[11] The tropical Pacific SST anomalies have perhaps the most variable hindcast skill over the 1960–2010 period (Figure 1e). MIROC5 has significant hindcast skill in 1960s, 1970s, 1980s, 1990s, and the oughts. HadCM3 shows highly variable behavior in the tropical Pacific, with high and significant hindcast skill in the 1970s and then significantly negative skill in 1980s, 1990s, and the oughts. Except for small but significant skill in 1980s and the oughts, the GFDL-CM2.1 does not show any significant

hindcast skill in the tropical Pacific. CCSM4 also does not show any significant hindcast skill in the tropical Pacific. The tropical Atlantic (Figure 1f) shows a more stable behavior with at least two ESMs showing significant hindcast skill in all decades except 1980s. The GFDL-CM2.1 shows significant skill in 1960s and 1970s, followed by MIROC5 in 1960s, 1990s, and the oughts, and HadCM3 in 1960s, 1970s, and 1990s. CCSM4 shows significant hindcast skill in 1960s, 1990s, and the oughts. The MME shows significant skill in 1960s, 1970s, 1990s, and the oughts. Overall, of the 30 combinations of averaged regions and decades shown in Figure 1, there is significant hindcast skill in 17 combinations in GFDL CM2.1, 25 combinations in MIROC5, 17 combinations in HadCM3, 9 combinations in CCSM4, and 25 combinations in MME. Thus, although model dependent, there is significant decadal SST hindcast skill in this subset of CMIP5 ESMs.

Table 2. Major, Low-Latitude Volcanic Eruptions From 1960 to 1991

Time	Volcano and Country	Latitude-Longitude	Remarks (VEI – Volcanic Explosivity Index)
February–May 1963	Mount Agung, Bali, Indonesia	8°20′31.2″S 115°30′28.8″E	VEI 5; ~1 km ³ material ejected
June–July 1968	Fernandina, Galapagos Island, Ecuador	0°22′S, 91°33′W	VEI 4; >0.1 km ³ material ejected
October–December 1974	Volcan de Fuego, Guatemala	14°28′58″N 90°52′58″W	VEI 4; >0.1 km ³ material ejected
March–April 1982	El Chichón, Chiapas, Mexico	17°21′36″N 93°13′40″W	VEI 5; ~1 km ³ material ejected
November 1985	Nevado del Ruiz, Colombia	4°53′43″N 75°19′21″W	VEI 3; > 10,000,000 m ³ material ejected
June 1991	Mount Pinatubo, Philippines	15°08′30″N 120°21′00″E	VEI 6; ~10 km ³ material ejected

5. Impact of Volcanic Eruptions on Decadal Hindcast Skill of Global and Regional SSTs

[12] Having found decadal fluctuations in decadal hindcast skills of all four ESMs and the MME, years and geographic

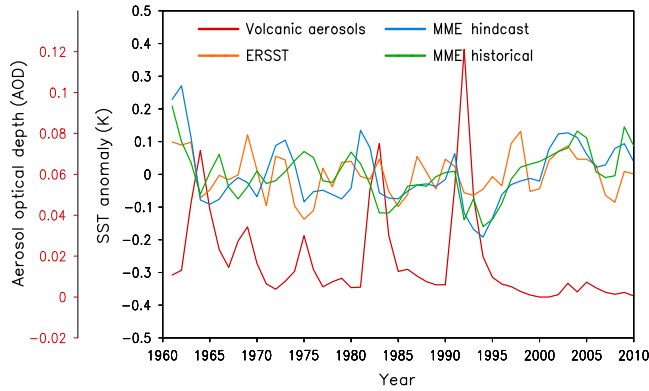


Figure 2. Time series of (red) annual average anomalous volcanic aerosol optical depth (1961–1999) and detrended global average SST anomalies (1961–2010) from (orange) ERSST, (blue) CMIP5 decadal hindcasts, and (green) CMIP5 historical experiments. The CMIP5 SST anomalies are multimodel ensemble (MME) averages.

locations of major volcanic eruptions were tabulated (Table 2) and compared with Figure 1. Also, time series of annual average AOD prescribed in CMIP5 experiments, global average SST anomalies from ERSST, the MME hindcast, and the MME historical simulation were plotted (Figure 2). A comparison of Table 2 with AOD in Figure 2 shows that the Mount Agung, Bali, Indonesia (1963); the Fernandina, Galapagos Island, Ecuador (1968); the Volcan de Fuego, Guatemala (1974); the El Chichón, Chiapas, Mexico (1982); and the Mount Pinatubo, Philippines (1991) eruptions are prominent in the AOD time series.

[13] Comparing the AOD time series with global average ERSST, MME historical SST, and MME hindcast SST anomalies in Figure 2, it is clear that the three global average SST anomalies decreased immediately or within a few months after the volcanic eruption-related AOD peaks in 1963, 1968, 1974, 1982, and 1991. Following the decreases, all three SST anomalies gradually recovered toward their pre-eruption values. The SST decreases following AOD peaks and subsequent increases in the three time series, however, were not in phase. It is clear that the relatively substantial and significant prediction skill of global SSTs in 1960s and 1990s is associated with impacts of volcanic eruptions, via changes in AOD in these four ESMs, on the ocean-atmosphere system and subsequent recovery of the system.

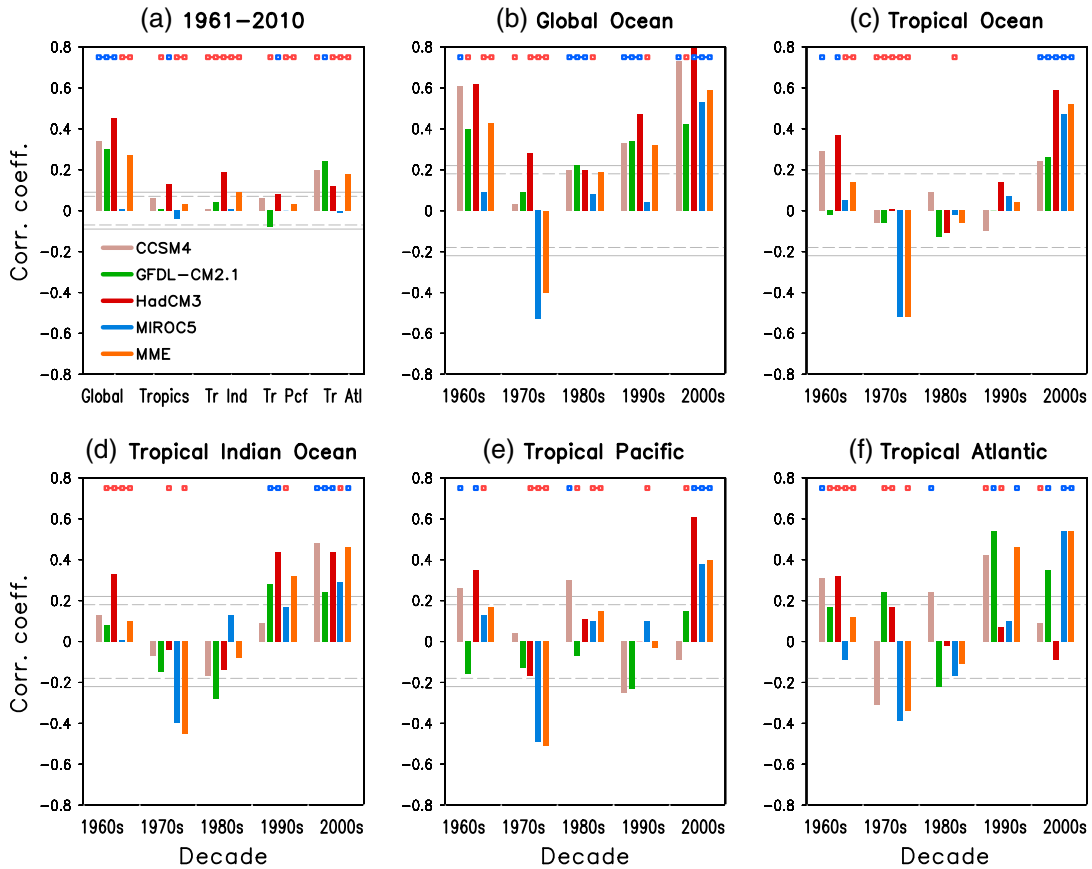


Figure 3. Same as Figure 1, but for SST from CMIP5 historical experiments. The correlation coefficients for HadCM3 during the last decade were based on data from 2001 to 2005. Red (blue) dots near the top of each panel denote that the correlation for SST from the decadal hindcast (historical) experiments is above the 95% significance level and is larger than the corresponding correlation with SST from the historical (decadal hindcast) experiments.

6. Impact of Model Initialization

[14] To test the hypothesis of influence of volcanic eruptions on skill further and to assess the impact of initialization of the ESM experiments with observed data, simulation skills were also estimated using the ESM historical simulations and compared with hindcast skills. Figure 3 shows the skills for the entire 1961–2010 period for various averaging regions and for each 10 year period in each averaging region. For the entire 1961–2000 period (Figure 3a), all models except MIROC5 show significant skill of global average and tropical Atlantic average SST anomalies. HadCM3 shows significant skill in all averaging regions, whereas MIROC5 has zero skill in all regions. The time dependence of decadal SST simulation skill in various averaging regions is shown in Figures 3b–3f.

[15] Analyses of the results shown in Figures 1 and 3 show that, of the 30 combinations of averaged regions and decades shown in each of the two figures, significant hindcast skill improves in 11 combinations in GFDL CM2.1, 21 combinations in MIROC5, 10 combinations in HadCM3, 4 combinations in CCSM4, and 16 combinations in MME when these models are initialized with observed data. This comparison of hindcast and simulation skills is shown in Figure 3 with blue (red) dots marking regions and decades in which simulation skill is significant and larger (smaller) than hindcast skill. As this comparison shows, the biggest impact of model initialization is on MIROC5 hindcasts as compared to uninitialized simulations. In the 2001–2010 decade, however, simulation skill was higher than hindcast skill in almost all combinations in all four ESMs and MME. These results, however, clearly show that the improvement in decadal hindcast skills of SST anomalies over ocean basin size averaging regions due to model initialization is model dependent.

7. Discussion and Conclusions

[16] The HadCM3, GFDL-CM2.1, CCSM4, and MIROC5 ESMs used in this study and the MME formed by combining the four ESMs' outputs show statistically significant decadal hindcast skill of global average and tropical ocean basin average SST anomalies during 1961–2010. The maximum correlation coefficients between simulated and observed SST anomalies over a decade, as an indicator of predictability skill, are approximately 0.65. The skill, however, varies by averaging region, decade, and model. The highest number (region and decade combinations) of significant decadal hindcast correlations are in MIROC5 and the smallest in CCSM4; in HadCM3 and CCSM4 hindcasts, there is some negative skill, especially in tropical Pacific basin. The largest number (regions) of significant skill in all four ESMs and the MME are in 1960s.

[17] It was also found that SSTs in all averaging regions decrease after a moderate to large volcanic eruption, with the maximum lagged correlation between specified AOD and SSTs approximately 1 year after eruption in both hindcast and historical experiments in all ESMs and the MME.

[18] It is clear that individual, moderate, and large volcanic eruptions influence SSTs globally and are one of the sources of decadal SST predictability. In the 2001–2010 decade, decadal SST hindcast skill was lower than simulation skill in all but five basin-decade combinations of 25 such

combinations. This result implies that model initialization worsened SST simulation skill compared to uninitialized historical experiments in the 2001–2010 decade. In the absence of detailed experiments to assess this difference in skills, we offer several speculations. Solomon *et al.* [2011] suggested that relatively small volcanic eruptions in this decade may have cooled SSTs enough to mitigate global warming. Another possibility is that, while the external forcing associated with volcanic aerosols was much smaller in the 2001–2010 decade compared to previous decades, the greenhouse gas forcing was much stronger. The latter may have degraded the influence of initialization in the hindcast experiments. Additionally, the global warming effect of greenhouse gases is stronger in the tropics than in the extratropics. This could lead to a larger simulation skill for the tropical SSTs. Also, deep ocean temperature observations are still sparse, and the ocean initialization in decadal hindcast experiments still depends largely on the ocean data assimilation system employed to initialize the ESMs, affecting decadal hindcast skill.

[19] While the results of this study are very interesting and encouraging in the quest for decadal climate predictability, it must be noted that the volcanic eruptions included in the specified AOD in the CMIP5 experiments would not normally be available for prediction experiments because volcanic eruptions themselves cannot be predicted at decadal time scale. What the results of this study show, however, is that after an eruption occurs, the ESMs used in this study appear to respond reasonably accurately to the injection of volcanic aerosols in the model atmosphere and the response of the SSTs then provides significant predictive skill for several years. A further study is in progress to analyze whether the SST hindcast results described here translate to skillful decadal hindcasts of land climate in the ESMs. Whatever may be the outcome of the land climate hindcast study, these early results of the CMIP5 decadal hindcast experiments appear to have begun well the quest for decadal climate predictability.

[20] **Acknowledgments.** This research is supported by the U.S. Department of Agriculture-National Institute of Food and Agriculture under grant 2011-67003-30213 in the NSF-USDA- DOE Earth System Modelling Program. The authors are grateful to Doug Smith (U.K. Meteorological Office-Hadley Centre, U.K.), Tom Delworth (NOAA-Geophysical Fluid Dynamics Laboratory, U.S.A.), and Toru Nozawa (National Institute for Environmental Studies, Japan) for discussions about the CMIP5 experiments conducted with their respective models and about the aerosol data used in these experiments. We thank the editor and two anonymous reviewers for their constructive comments that have substantially improved this paper.

References

- Ammann, C., G. Meehl, W. Washington, and C. Zender (2003), A monthly and latitudinally varying forcing data set in simulations of 20th century climate, *Geophys. Res. Lett.*, *30*(12), 1657, doi:10.1029/2003GL016875.
- Hansen, J. E., et al. (2002), Climate forcing in Goddard Institute for Space Studies S12000 simulations, *J. Geophys. Res.*, *107*(D18), 4347, doi:10.1029/2001JD001143.
- Keenlyside, N., M. Latif, J. Jungclauss, L. Kornbluh, and E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, *453*, 84–88, doi:10.1038/nature06921.
- Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran (2000), Multimodel ensemble forecasts for weather and seasonal climate, *J. Clim.*, *13*, 4196–4216.
- Meehl, G. A., et al. (2009), Decadal prediction: Can it be skillful?, *Bull. Am. Meteorol. Soc.*, *90*, 1467, doi:10.1175/2009BAMS2778.1.
- Mehta, V. M., C. L. Knutson, N. J. Rosenberg, J. Rolf Olsen, N. A. Wall, T. K. Bernadt, and M. J. Hayes (2013), Decadal climate information needs of stakeholders for decision support in water and agriculture

- production sectors: A case study in the Missouri River Basin, *Weather Clim. Soc.*, 5, 27–42, doi:10.1175/WCAS-D-11-00063.1.
- Pohlmann, H., J. H. Jungclaus, A. Kohl, D. Stammer, and J. Marotzke (2009), Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic, *J. Clim.*, 22, 3926–3938, doi:10.1175/2009JCLI2535.1.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang (2002), An improved in situ and satellite SST analysis for climate, *J. Clim.*, 15, 1609–1625.
- Sato, M., J. Hansen, M. P. McCormick, and J. Pollack (1993), Stratospheric aerosol optical depth, 1850–1990, *J. Geophys. Res.*, 98, 22,987–22,994.
- Smith, D., S. Cusack, A. Colman, A. Folland, G. Harris, and J. Murphy (2007), Improved surface temperature prediction for the coming decade from a global circulation model, *Science*, 317, 796–799, doi:10.1126/science.1139540.
- Solomon, S., J. S. Daniel, R. R. Neely III, J.-P. Vernier, E. G. Dutton, and L. W. Thomason (2011), the persistently variable “background” stratospheric aerosol layer and global climate change, *Science*, 333, 866–870, doi:10.1126/science.1206027.
- Stenchikov, G., K. Hamilton, R. J. Stouffer, A. Robock, V. Ramaswamy, B. Santer, and H.-F. Graf (2006), Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models, *J. Geophys. Res.*, 111, D07107, doi:10.1029/2005JD006286.
- Tatebe, H., et al. (2012), Initialization of the climate model MIROC for decadal prediction with hydrographic data assimilation. JMSJ Special issue on the recent development on climate models and future climate projections, *J. Met. Soc. Japan*, 90A, 275–294, doi:10.2151/jmsj.2012-A14.
- Wilks, D. S. (1995), *Statistical Methods in the Atmospheric Sciences*, 467 pp., Academic Press, Oxford, U.K.
- Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng (2012), A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content, *J. Clim.*, 25, 5173–5189, doi:10.1175/JCLI-D-11-00595.1.
- Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg (2007), System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies, *Mon. Weather Rev.*, 135, 3541–3564, doi:10.1175/MWR3466.1.