Characteristics of tropical cyclones in high-resolution models in the present climate

Daniel A. Shaevitz,[1] Suzana J. Camargo,[2] Adam H. Sobel,[1,2,3] Jeffrey A. Jonas,[4,5] Daehyun Kim,[2,6] Arun Kumar,[7] Timothy E. LaRow,[8] Young-Kwon Lim,[9,10] Hiroyuki Murakami,[11] Kevin Reed,[12] Malcolm J. Roberts,[13] Enrico Scoccimarro,[14,15] Pier Luigi Vidale,[16] Hui Wang,[7] Michael F. Wehner,[17,18] Ming Zhao,[19] and Naomi Henderson[2]

[1] Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA.

[2] Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA.

[3] Department of Earth and Environmental Sciences, Columbia University, New York, NY, USA.

[4] Center for Climate System Research, Columbia University, New York, NY, USA.

[5] NASA Goddard Institute for Space Studies, New York, NY, USA.

[6] Current address: Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA.

[7] NOAA/NWS/NCEP Climate Prediction Center, College Park, MD, USA.

[8] Center for Ocean Atmospheric Prediction Studies, Florida State University, Tallahasee, FL, USA.

[9] NASA Goddard Space Flight Center, GMAO, GreenBelt, MD, USA.

[10] Goddard Earth Sciences Technology and Research/I.M. Systems Group, Greenbelt, MD, USA.

[11] International Pacific Research Center, University of Hawaii at Manoa, Honolulu, HI, USA.

[12] National Center for Atmospheric Research, Boulder, CO, USA.

[13] Met Office, Hadley Center, Devon, United Kingdom.

[14] Istituto Nazionale di Geofisica e Vulcanologia (INGV), Bologna, Italy.

[15] Centro Euro-Mediterraneosui Cambiamenti Climatici (CMCC), Bologna, Italy.

[16] NCAS-Climate, University of Reading, Reading, United Kingdom.

[17] Lawrence Berkeley National Laboratory,Berkeley, CA, USA.

[18] University of California, Berkeley, Berkeley, CA, USA.

[19] NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA.

Corresponding author: Suzana J. Camargo, Lamont-Doherty Earth Observatory, Columbia University, 61 Rt. 9W, Palisades, NY 10964-8000, USA. suzana@ldeo.columbia.edu

Abstract. The global characteristics of tropical cyclones (TCs) simulated by several climate models are analyzed and compared with observations. The global climate models were forced by the same sea surface temperature (SST) fields in two types of experiments, using climatological SST and interannually varying SST. TC tracks and intensities are derived from each model's output fields by the group who ran that model, using their own preferred tracking scheme; the study considers the combination of model and tracking scheme as a single modeling system, and compares the properties derived from the different systems. Overall, the observed geographic distribution of global TC frequency was reasonably well reproduced. As expected, with the exception of one model, intensities of the simulated TC were lower than in observations, to a degree that varies considerably across models.

1. Introduction

The impact of tropical cyclones (TCs) on society makes it important to understand how their characteristics might change in the future. Global climate models, also known as General Circulation Models (GCMs), are important tools for studying this problem. In a GCM, one has the ability to simulate the climate organically; if the model has sufficient resolution and physics to provide a plausible simulation of TCs as well, then one can use the model to examine how climate controls the statistical properties of TCs. One can explore, in particular, the behavior of TCs under different climate scenarios.

Many studies (e.g., *Manabe et al.* 1970; *Bengtsson et al.* 1982; *Vitart et al.* 1997; *Camargo et al.* 2005) have shown that GCMs, even at relatively low resolution, are capable of generating storms that have similar characteristics as observed TCs. More recently, studies that have used higher resolution atmospheric GCMs forced with prescribed sea surface temperatures (SSTs) (e.g., *Bengtsson et al.* 2007a; *LaRow et al.* 2008; *Zhao et al.* 2009) have demonstrated these high-resolution models' remarkable ability to simulate realistic distributions of TCs.

In order to use GCMs for projections of possible future changes in TC activity, it is necessary to assess their ability to reproduce the characteristics of observed TCs in the present climate. These characteristics include the climatological spatial, temporal, and intensity distributions as well as the interannual variability of TCs. This work is an intercomparison of the ability of 9 high-resolution GCMs to simulate TCs. The models have resolutions that vary from 28 to 130 km, with different parameterizations and dynamics.

Two of the models have done simulations at multiple resolutions, while a single resolution is available for our analysis of the other models.

The simulations analyzed were performed for the U.S. CLIVAR Hurricane Working Group. The objective of this working group was to have a better understanding of the differences among high-resolution models in simulating TC activity, in the present climate as well as in warmer climate scenarios. In order to do that, a set of common experiments with the same forcings and prescribed SST was performed by all modeling groups. Here we analyze the characteristics of TC activity in the simulations produced by the working group over SST distributions derived from observations taken in the late 20th century (1981-2005 for the climatology simulations and 1981-2009 for the interannual simulations).

Observed TC tracks and intensities are derived from atmospheric measurements — in situ and remote — by human forecasters. With climate models, it is necessary to apply objective tracking schemes to the model output fields to obtain the tracks and intensities. The criteria applied to the models can be different from those applied to observations; a model storm is not necessarily required to meet the same thresholds for intensity as an observed one would be in order to be classified as a TC. It has been found that when allowance is made for the fact that model TCs are weaker and larger than those observed, the resulting spatio-temporal distributions of TC tracks resemble those observed enough to be useful — for example, in seasonal forecasting — even in quite low-resolution models [*Camargo and Barnston*, 2009; *Camargo et al.*, 2010].

In the present study, we examine the TCs derived from each model's output by the group who ran that particular model, using their own preferred tracking scheme. We consider the combination of model and tracking scheme to be a "modeling system" and compare the outputs from each system. In the interests of brevity, we will refer to these modeling systems below simply as "models", taking the tracking scheme as implicit, though our expectations about the sensitivities of the results to tracking schemes are discussed in several points.

This approach implicitly makes allowances for the different resolutions and physics of each model, resulting in different TC intensities. It is consistent with the way each model has been used in previous single-model studies. Using each group's own tracks also allows each model to be seen as each group intended, to the extent that tracking schemes have tunable parameters whose adjustment can allow some gross aspects of the statistics to be brought closer to those observed. An alternative approach that could be considered is to use a bias correction procedure to obtain values closer to observations as was done in *Murakami et al.* [2012] for TC frequency and in *Zhao and Held* [2010] for TC intensity. We will leave the bias correction analysis of TC activity for future work.

It is also of interest to compare the different models using the same tracking scheme, so that the differences in results are purely attributable to the differences in the models themselves. This has been done by *Horn et al.* [2014], who also used multiple tracking schemes to study the sensitivity of the analysis to the tracking scheme used. We focus our analysis on the following TC characteristics: TC frequency, intensity and lifetime. Other TC characteristics could potentially be explored in these models, such as TC size, which only recently has been receiving more attention in observations [*Dean et al.*, 2009; *Chavas and Emanuel*, 2010; *Knaff et al.*, 2014] and idealized models [*Chavas and Emanuel*, 2014]. Analysis of the rainfall associated with TCs in a subset of the models considered here was presented in *Villarini et al.* [2014] and *Scoccimarro et al.* [2014].

This paper is organized as follows. The data, models, and experiments are discussed in section 2. Results from the climatological and historical forced simulations are described in section 3. Finally, conclusions are given in section 3.2.

2. Models and Data

The data used for this study consists of TC tracks from nine GCMs. The models were forced with two different SST boundary conditions, monthly climatologically averaged (seasonally varying) SSTs and monthly interannually varying SSTs. The SSTs were obtained from the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set [*Rayner et al.*, 2003] and the climatological SST was obtained by averaging the monthly data over the period 1981-2005, which was previously used in *Held and Zhao* [2011]. The number of years in the climatological simulations performed by each group varied from 5 years to 20 years, as shown in Table 2.

Each group used the output of their simulations to detect and track the model TCs, using their own tracking algorithm. Tracks for these TCs were generated and their characteristics were analyzed here. The sensitivity of the models to the different tracking schemes is currently being analyzed by members of the working group.

Outputs from nine GCMs were analyzed in this study, as summarized in Table 1, namely: Community Atmosphere Model version 5.1, or CAM5.1 [*Wehner et al.*, 2014]; European Centre for Medium-Range Weather Forecasting - Hamburg, or ECHAM5 [*Roeckner et al.*, 2003; *Scoccimarro et al.*, 2011]; Florida State University, or FSU [*LaRow et al.*, 2008]; NASA Goddard Earth Observing System Model version 5, or GEOS-5 [*Rienecker et al.*, 2008]; National Centers for Environmental Prediction Global Forecasting System, or GFS [*Saha et al.*, 2014]; NASA Goddard Institute for Space Studies, or GISS [*Schmidt et al.*, 2014]; Met Office Hadley Centre Model version 3 - Global Atmosphere 3.0 (GA3) configuration, or HadGEM3 [*Walters et al.*, 2011]; Geophysical Fluid Dynamics Laboratory High Resolution Atmosphere Model, or HiRAM [*Zhao et al.*, 2009]; and Meteorological Research Institute, or MRI [*Mizuta et al.*, 2012; *Murakami et al.*, 2012]. The model resolutions vary from 28 to 130 km.

The models have different tracking schemes, most of them with very similar characteristics, based on the original tracking schemes in *Bengtsson et al.* [1982] and *Vitart et al.* [2007]. These tracking schemes look for vortices with a minimum of sea level pressure, a maximum of low-level vorticity and a warm core [*Camargo and Zebiak*, 2002; *Walsh*, 1997; *Vitart et al.*, 2003; *Zhao et al.*, 2009; *Murakami et al.*, 2012]. The main difference among the schemes is how they define the warm core and the thresholds used to define the model TC. An exception is the HadGEM3, which uses a tracking scheme originally developed for extra-tropical (cold core) cyclones [*Hodges*, 1995] and modified to track warm core vortices [*Bengtsson et al.*, 2007a; *Strachan et al.*, 2013]. More detailed descriptions of the tracking schemes are given in the Appendix.

We compare the model TC characteristics with the observed TC data. For the North Atlantic and eastern and central North Pacific the best-track datasets from the National Hurricane Center are used [*Landsea and Franklin*, 2013; *NHC*, 2013]. In the case of the western North Pacific, North Indian Ocean and southern hemisphere, the TC data is from the best-track datasets from the Joint Typhoon Warning Center [*Chu et al.*, 2002; *JTWC*, 2014].

3. Results

3.1. Climatology

We first examine the climatological simulations, in which all models are forced with the same monthly, climatological, seasonally varying SST fields. As there is no interannual variability in these SST fields, we can use them to assess the level of internal, unforced variability in the models' TC activity. We will also compare the mean TC activity in each model with the observations, globally and in different basins.

3.1.1. TC Frequency

There are on average approximately 80 TCs observed every year across the globe [*Emanuel*, 2003]. Figure 1 shows the distribution of the number of TCs per year for all models along with the observations. There are large differences in the number of TCs between the different models. Different models run at approximately the same resolution do not have similar mean numbers of TCs (e.g., the LR CAM5.1, FSU, GFS, and GISS models all have resolutions of roughly 100 km, but the mean number of TCs per year varies from about 10 to over 100.)

At the same time, the absolute number of TCs in each model is somewhat dependent on the tracking scheme applied; higher thresholds result in fewer TCs. This is particularly evident in the CAM5.1 models, where the same thresholds were used for both the low resolution and high resolution simulations, resulting in a very low number of TCs in the LR CAM5.1 model. Application of strictly uniform tracking schemes, with no allowance for the different intensities in different models (whether due to resolution or other factors) would almost certainly produce even larger differences in the total numbers of TCs from model to model. By using each group's own tracking scheme, we allow some compensation for the different TC intensities, in order to allow more productive comparison between other aspects of the results, such as the spatial and seasonal distributions of TC genesis and tracks, in the way that they would be shown in single-model studies by the individual groups.

The three resolutions of the HadGEM3 model show an increase in the number of TCs with increasing resolution, though it does not increase linearly. The tracking algorithm for all resolutions of the HadGEM3 model use the same threshold for the 850-hPa relative vorticity after being filtered to a standard spectral resolution of T42 as described in *Strachan et al.* [2013]. Thus, the increase in the number of TCs with increasing resolution is not an artifact of the tracking scheme.

Figure 2 shows the mean number of TCs formed per year in each ocean basin. The total number of TCs formed in each basin per year is shown at the top of the figure and the percentage of all TCs that formed in each basin is shown at the bottom. Due to the large differences in the total numbers of global TCs reported by each model, it is more illustrative to compare the percentages of the TCs that form in each basin, rather than the total number of TCs, to the observations.

There are clear differences among the models in the distribution of TCs across basins, particularly in the North Atlantic and Pacific. Several of the models (ECHAM5, GISS, and all resolutions of the HadGEM) have percentages much lower than that observed in the North Atlantic. Three of the models (ECHAM5, FSU, and GISS) have a significantly lower percentage than that observed in the Eastern North Pacific, while the CAM5.1 (at both resolutions) and GFS have a much higher percentage than observed in the Eastern North Pacific. In the Western North Pacific, the CAM5.1 models have smaller percentages than observed, and the ECHAM5 and GISS models have larger percentages than observed.

This is consistent with previous studies that have found that low-resolution models tend to have a large percentage of TCs in the Western North Pacific and very few TCs in the North Atlantic [*Camargo et al.*, 2005; *Camargo*, 2013]. Also of note is that the discrepancies in the partitioning between the Western and Eastern North Pacific in the CAM5.1 models are partially linked to a bias toward too many TCs in the Central North Pacific.

One interesting observation is that there are larger differences in TC distributions between one model and another, than between versions of the same model at different horizontal resolutions. The TC distributions obtained by the versions of the CAM5.1 with different resolutions are very similar, and the same is true of the HadGEM3 models. This suggests that the global and regional distributions of TCs is mainly determined by the characteristics of the models (e.g., parameterizations, convection scheme), with model resolution not being as important. While the tracking schemes are also different, our expectation is that the usage of different tracking schemes reduces the apparent differences between models, particularly in overall TC frequency. As will be seen below, the intensities of the simulated TCs are quite different in different models, and the different thresholds in the tracking schemes adjust for this to a large degree. If the same tracking scheme (including the specific thresholds) used to detect TCs in HiRAM were applied to the GISS model, for example, very few TCs would be detected.

In order to study the geographic patterns of TC occurrence, we will use track density, defined as the number of TCs that pass through a 5° x 5° box per year. Figure shows the track density for all models and observations. The observed track density shows a region of very high density off the western coast of Central America and the eastern coast of Asia, along with regions of high density in the North Atlantic, South Indian, and off the eastern coasts of Australia and India.

Consistent with the basin averages, the models have different patterns of track density. The GISS model has a similar pattern to the observations, with some key differences. The most striking difference is the lack of a region of high track density off the western coast of Central America, which is notoriously difficult to simulate with lower resolution GCMs [*Camargo et al.*, 2005]. Other differences include a higher density around India, the region of high density off the eastern coast of Asia extending further to the east, and a lower density in the North Atlantic. The HiRAM model has a remarkably similar pattern to the observations globally. The FSU model has higher density in the North Atlantic and South Indian along with lower density off the eastern coast of Central America. The ECHAM5 model has very low density in the North Atlantic and South Indian, but similar density patterns to the observations in the Western Pacific and South Pacific. The ECHAM5 model also has a localized region of very high density directly on the eastern coast of India. The high resolution CAM5.1 model has a region of very high density off the western coast of Central America that extends too far westward and has much lower density off the eastern coast of Asia than the observations. The low resolution HadGEM3 model has small regions of high density in the correct locations. The higher resolution HadGEM3 models have higher density in these regions, which expand covering larger areas. The global mean densities in the low resolution CAM5.1 and GFS models are much lower than observed. Also shown is the multi-model mean (MMM) track density (using only the high-resolution version of the CAM5.1 and HadGEM3 models). The MMM pattern's similarity to the pattern in observations is greater than those in many of the individual models, but the magnitudes of the maxima are not as high as in observations.

These results are consistent with the findings of *Strazzo et al.* [2013] which examined track densities of the FSU and HiRAM simulations. *Strazzo et al.* [2013] showed that the HiRAM density distribution is very similar to the observed distribution, while FSU model has a higher density in the North Atlantic than is found in observations.

In addition to track density, it is useful to study where the simulated TCs form, or genesis density. Figure 4 shows the genesis density of all the models and observations. Genesis density is defined as the number of TCs that form in a 5° x 5° box per year. The overall differences in the patterns of the genesis density between the models and observations are similar to the differences in the track density described above. Consistent with the observations, all the models have narrower meridional bands of high genesis density as compared to track density. This occurs because the TCs tend to form in low-latitudes and travel poleward, causing the track density to have a greater meridional spread than the genesis density. Similarly to the case of track density, the genesis density MMM pattern is closer to the observations than many of the individual models.

It can be easier to distinguish patterns in the distributions by examining certain spatial or temporal dimensions. Fig.5(a) shows the genesis as a function of latitude of each model and the observations. For the CAM5.1 and HadGEM3 models, only the highest resolution simulations are shown. The observations have a large peak at 10° north, a smaller peak at 10° south, and no TC formation directly at the equator. All of the models have peaks at roughly the same latitudes as the observations, with the FSU and GEOS-5 model having a peaks closer to the equator, especially in the southern hemisphere, and the ECHAM5 model, having peaks poleward than the observations. In addition, the FSU model has a high number of storms forming very near the Equator in the southern hemisphere. The ECHAM5 model's southern hemisphere peak has a fatter tail and has non-zero genesis extending to higher latitudes than the observations and all other models. Although the GFS model has fewer TCs than in observations, the maxima in genesis location occur at roughly the same latitudes and with similar relative magnitude as the observations.

Fig.5(c) shows the genesis as a function of longitude for the models and observations. The observations have two sharp peaks at roughly 90E and 110W (corresponding to the maxima in the South Indian and western coast of Central America in Fig. 4), a broader peak at roughly 150E (corresponding to the maxima off the eastern coast of Asia in Fig. 4), and near-zero genesis near and east of the dateline. Three of the models (GISS, FSU, and ECHAM5) have much lower Central American 110W peak than the observations, with the GISS model producing virtually no TCs. The FSU model has peaks at 55° (off the eastern coast of Africa) and 50W (North Atlantic) that are not present in any other model or the observations. The ECHAM5 model has a very strong peak at 85E (off the eastern coast of India). The HiRAM model exhibits a pattern remarkably similar to the observations.

Another metric of interest is the seasonal cycle of TC formation. Fig. 5(b) shows global genesis as a function of month for models and observations. The observations show a fairly smooth seasonal cycle with a clear maximum between August and September and a minimum around April. In general, the models have a significantly weaker seasonal cycle than the observations, i.e. the difference between the number of TCs in the second half of the year and the first half of the year is less than the same difference in the observations.

The TC seasonal cycles in different basins are shown in Fig. 6[1]. The basins in the northern hemisphere typically have a broad peak in the second half of the year and few TCs in the first half of the year, with exception of the North Indian Ocean. In the Western North Pacific, the GISS, HiRAM, FSU, HR HadGEM3, and ECHAM5 models are able to roughly reproduce the peak in the second half of the year, while the other models have no peak. In the Eastern North Pacific, the HiRAM3, HR HadGEM3, HR CAM5.1, and GFS models are able to reproduce the August peak while the other models have very low density throughout the year in this basin. However, HR CAM5.1 has a second peak in October and November that does not occur in the observations. In the North Atlantic, the HiRAM3, FSU, HR CAM5.1, and GFS models reproduce the second half of the year peak. Also of note is that the FSU model has a peak in the Western North Pacific that is roughly three months later than in observations, while it has a peak in the North Atlantic roughly one month earlier than observed. Most models are able to capture the bimodal distribution in the North Indian Ocean, with exception of the ECHAM5. All models are able to reproduce the observed peak in the early part of the year in the South Pacific and Australian basins. In contrast, in the South Indian basin, the CAM5.1 and FSU models have the wrong seasonality with a peak in the second half of the year.

3.1.2. TC Intensity

Along with the frequency of TCs, it is important to examine TC intensity. Although the global climate models here are considered “high-resolution”, it is not expected that they would be able to reproduce the most intense TCs (category 4 and 5 hurricanes), which would require even higher resolution to be able to simulate those intensities (see e.g., *Bender et al.* [2010]). A significant fraction of the models has TCs that reach category 4, but only one model has TCs that reach category 5.

The accumulated cyclone energy (ACE) of a TC is the sum of the squares of the TC's maximum wind speed, summed over all 6-hour intervals in which the maximum wind speed is at least tropical storm strength (35 kt). Adding the ACE of individual TCs can produce a total ACE for a spatial or temporal region, e.g., a basin ACE or a seasonal ACE. Thus, a larger value of total ACE could correspond to stronger TCs, more TCs, and/or TCs that last longer. Figure 7 shows the total ACE (averaged per year) for each basin. The top panel shows the total ACE of each basin and the bottom panel shows the percentage of the global ACE that occur in each basin. The observations have large values of ACE in the Western North Pacific (40%) , followed by the eastern North Pacific, North Atlantic and South Indian Ocean (15%), with the Australian and South Pacific contributing with about 5% of the global ACE and a very low value of ACE in the North Indian Ocean. All models are able to reproduce the large ACE percentage in the Western North Pacific, with the ECHAM5 and FSU models having a very low ACE percentage in the Eastern North Pacific. The ECHAM5 and GISS models have a relatively large ACE percentage in the South Pacific, while the HadGEM3 models (all resolutions) have an anomalously high ACE percentage in the South Indian Ocean.

The top panel of Fig. 8 shows the distribution of the maximum wind speed achieved by each TC in all models and the observations. The vertical lines represent boundaries of the Saffir-Simpson hurricane intensity scale [*Saffir*, 1977]. The models seem to separate into four regimes of intensities. The HR CAM5.1 has an intensity distribution similar to observations, with a significant number of category 2 hurricanes and even the ability to produce the most intense TCs, i.e. categories 4 and 5 storms. The HiRAM, FSU, and HR HadGEM3 models have many tropical storms and category 1 TCs and some category 2 TCs. The ECHAM5, GEOS-5, and GFS models have mostly tropical storms. The GISS model's TCs are almost all of tropical depression intensity, with only a very small number of weak tropical storms. The difference between the intensity distributions among the models cannot simply be a result of the models' different resolutions. For example, the GEOS-5 model has a horizontal resolution similar to the HiRAM model, but has significantly weaker TCs. On the other hand, the FSU model has some of the strongest TCs, but does not have one of the highest resolutions among the models.

In order to better understand the effect of model resolution on simulated TC intensities, it is instructive to examine the differences in the intensity distributions of the same models run with multiple horizontal resolutions. Histograms of the maximum wind speeds for the CAM5.1 and HadGEM3 models using various different resolutions are shown in the bottom panels of Fig. 8. As expected, both the CAM5.1 and HadGEM3 models show an increase in the mean TC intensity with higher resolution. The increase in intensity of the HR HadGEM3 and HR CAM5.1 models can be also seen as an elongation of the tails of the distributions into higher TC categories.

3.1.3. TC Lifetime

TC lifetime distributions in models and observations are shown in Fig. 9, with the TC lifetime histograms of the CAM5.1 and HadGEM3 models in different resolution given in the two bottom panels. There is a large variation in the TC lifetime among the models. The ECHAM5, GISS, and HR HadGEM3 models have TCs lasting longer than 40 days, while the GFS model has very few TCs lasting more than 10 days. This is most likely due to the different tracking schemes used, as they consider different criteria for when to form and end a TC. Of particular note is that for the models with simulations in multiple resolutions, the TCs in the higher resolution simulations have a slightly longer average duration than in the low-resolution ones. This is probably also an artifact of the tracking scheme, as if the same intensity thresholds are used for high-resolution simulations, which generate more intense storms, this will lead to longer-living storms.

3.2. Interannual Variability

In the previous section, we analyzed the model simulations forced with climatological SSTs, which characterizes the typical TC properties in the models, but does not simulate the TC interannual variability. Well known modes of climate variability in the atmosphere and ocean, most notably the El Niño-Southern Oscillation (ENSO), have been shown to affect the frequency and characteristics of TCs [*Camargo et al.*, 2010; *Iizuka and Matsuura*, 2008; *Bell et al.*, 2013]. In order to evaluate the ability of the models to accurately simulate the interannual variability of TCs, the models were also run while forced with historical monthly varying SST, as opposed to climatological mean SSTs. The number of ensemble members and years of the simulations are shown in Table 3.

Figure 10 shows the total number of TCs globally per year for the models and observations (top panel), as well as for the Western North Pacific, Eastern North Pacific, and North Atlantic, separately[2,3]. The global number of TCs in the models is similar to the observed numbers in all the models, but the global interannual variability is not well captured by the models. The three individual basins shown here present a greater similarity between the observations and model results, with the exception of the GISS model which has very few TCs in the North Atlantic and Eastern North Pacific and the FSU model which has very few TCs in the Eastern North Pacific.

In order to quantify the ability of the models to reproduce the interannual variability of observed TCs in different basins, we calculate the correlation coefficients between the model-simulated and observed ACE per year in each basin for each model in Table 4. Since the GISS model's TCs have very weak intensities that seldom exceed the ACE threshold of 35 kt, we define another metric, the model-ACE (MACE), as the sum of the squares of the TC's maximum wind speed, sampled at 6-hourly intervals without any intensity threshold (as was done in *Camargo et al.* [2005] for low-resolution models). The correlations of the models' yearly MACE in each basin with the yearly ACE of the observations are also shown in Table 4. The correlations in the North Atlantic and Pacific basins are much higher than the other basins. In particular, the FSU and HiRAM models have a correlation coefficient of 0.7 in the North Atlantic and the GEOS-5 model has a correlation coefficient of 0.7 in the Western North Pacific basin. Similar result are obtained when calculating the correlation of the number of TCs per year globally and per basin (shown in Table 6), the highest and significant values of the correlations occur in the North Atlantic for all models and in other basins (eastern and western North Pacific) only for the HIRAM model.

Figure 11 shows the differences in genesis density between composites of El Niño and La Niña years. The seasons for the El Niño and La Niña composites are defined separately for the northern and southern Hemispheres in Table 5[4]. The observations have a larger and stronger peak in genesis density off of the western coast of Central America in El Niño months than La Niña months. As the GISS and FSU models have very few TCs in this region, they are unable to reproduce this difference, while the HiRAM and GEOS-5 models are able to reproduce the difference.

A well known impact of ENSO on TC development is for average formation location to shift to the south and east in the Western North Pacific and to shift to the south and west in the Eastern North Pacific during El Niño years [*Chia and Ropelewski*, 2002]. Figure 12 shows the mean position of TC formation in the Western and Eastern North Pacific in La Niña and El Niño years. In the Western North Pacific, the models are able to reproduce the southeast shift in El Niño years, with exception of the FSU model which has an eastern shift, with no meridional change. In the Eastern North Pacific, all the models are able to simulate the southwest shift in El Niño years.

This work has described an intercomparison of several high-resolution atmospheric models of the present climate, forced with both climatological and historical SSTs, in their ability to simulate the characteristics of TCs seen in observations. Model TCs were compared to observational TCs in terms of frequency as well as spatial, temporal, and intensity distributions. A range of tracking schemes were applied by each individual group to derive TC tracks and intensities for all models, consistent with the way in which results from these models have been shown previously in single-model studies.

Overall the models were able to reproduce the geographic distribution of TC track density in the observations, with the HiRAM model, in particular, demonstrating the most similarity to observations. TC formation off the western coast of Central America was the most difficult region to correctly simulate, with the HiRAM, HR CAM5.1, and HadGEM3 models demonstrating superior performance.

The models tend to have a weaker seasonal cycle in this region than is found in observations, as some of the models are too active in the southern hemisphere basins (e.g., FSU in the South Indian Ocean, ECHAM5 in the Australian and South Pacific, GEOS-5 in the South Pacific) in the first half of the year. The models reproduce the observational seasonal cycle to varying degrees in each basin, with the HiRAM model showing arguably the best match to observations overall.

There is a wide range in TC intensities between the different models. Some, but not all, of this difference can be seen as a consequence of resolution, with higher resolution models being able to simulate stronger TCs. This effect can be most readily seen in the CAM5.1 and HadGEM3 models which were run at multiple resolutions.

Many previous studies have predicted a decrease in TC frequency and an increase in TC intensity in a warmer climate [*Knutson et al.*, 2010]. The prediction of a decrease in TC frequency is mainly from modeling studies, where GCM simulations of a warmer climate produce fewer TCs than the present climate, with a few notable exceptions (e.g., *Emanuel* 2013). Although some of the current models still have biases in reproducing the mean global number of TCs, they are able to reproduce other characteristics of the TC activity. These biases could be potentially corrected using statistical methods as was done in *Zhao and Held* [2010] and *Murakami et al.* [2012]. On the other hand, some of the models (especially HiRAM) are able to simulate the TC climatology remarkably well. It is particularly encouraging that in the simulations forced with historical SSTs, the models were able to reproduce the interannual variability of TC frequency in the North Pacific and Atlantic basins, with the HiRAM and GEOS-5 models showing particularly high correlation with observations in those basins. All models were also able to reproduce the general geographic shift in TC formation location during El Niño and La Niña years.

Appendix A: Tracking Schemes

Here we give a description of the tracking schemes used by the various modeling groups. In general the tracking schemes look for features in which there is a minimum of sea level pressure, a maximum in vorticity and the existence of a warm core. The schemes vary in the definition of the thresholds for the different variables and in the definition of the warm core, but all tracking schemes have similar characteristics that can be traced back to the original papers of *Bengtsson et al.* [1982] and *Vitart* [1998].

The GFDL tracking scheme [*Vitart*, 1998; *Zhao et al.*, 2009] was used to track storms in the HiRAM, GFS, and CAM5.1 models. In the case of the CAM5.1 it was modified to run on a highly parallel systems [*Prabhat et al.*, 2012]. The original Vitart scheme was used in the FSU and GEOS-5 models, while for the ECHAM5 model, the Vitart scheme was modified by the Walsh wind speed resolution dependent thresholds [*Walsh*, 1997]. The GFDL tracking scheme identifies TCs by locating grid points that have a relative vorticity maxima exceeding $3.5 \times 10^{-5} s^{-1}$ within a 6 degrees latitude x longitude box; a local minimum of sea level pressure within 2 degrees of the vorticity maximum and a local maximum anomaly in temperature between 300 and 500hPa, at least 1°C warmer than the surrounding environment, withing 2 degrees of the sea level pressure maximum. The resulting points are combined into trajectories by associating the closest successive detections within 400km of each other. The tracks are required to last at least 3 days and have a maximum suraface wind speed greater than 12 m/s during at least 2 days (not necessarily consecutive).

The GISS model used the *Camargo and Zebiak* [2002] detection scheme. This scheme, derived originally for seasonal forecasting using low-resolution models, is similar to the others in most respects, but obtains model-dependent thresholds by analyzing the tails of the probability distribution functions of specific variables found in each model's output: 850hPa vorticity, anomalous integrated temperature (850 to 300hPa), surface wind speed. The algorithm then finds grid points in which these variables are higher than the model-dependent thresholds and where there is a local minimum in sea level pressure, a positive local temperature anomaly (850 to 300hPa), a larger temperature anomaly in 850 hPa than in 300 hPa and higher wind speeds in 850hPa than in 300hPa. These points are then joined into tracks if they occur within 5 degrees of each other. Only tracks that last at least 1.5 days are considered. These tracks are then extended forwards and backwards in time by tracking a vorticity maximum which is above a relaxed vorticity threshold.

The MRI models tracking scheme is described in *Murakami et al.* [2012], six criteria are considered: (i) a maximum relative vorticity above $8 \times 10^{-5} s^{-1}$, (ii) maximum wind speed at 850hPa larger than 13m/s, (iii) sum of temperatures at 300, 500, and 700hPa higher than 0.8K, (iv) maximum wind speed at 850hPa is higher than at 300hPa, (v) in the North Indian Ocean only, the radius of the maximum mean wind speed must be less than 200 km from the storm center, (vi) the storm last at least 36 hours. If the storm satisfies the criteria intermittently, multiple storms are considered, only one single time-step failure is allowed.

The HadGEM3 model tracking scheme is based on the Hodges method [*Hodges*, 1995, 1996, 1999] developed originally to track extra-tropical cyclones. The application of the Hodges method to tropical cyclones is described in *Bengtsson et al.* [2007a] and *Stra-*

chan et al. [2013], where the warm core criteria was refined. The 850hPa relative vorticity is used on a spectral resolution of T42, making this method resolution independent. All vorticity centers with intensity greater than $0.5 \times 10^{-5} s^{-1}$ at T42 are tracked, if they last at least 2 days then they are further analyzed. The 850hPa vorticity is then applied on a finer resolution (T63), and must reach at least a value of $6 \times 10^{-5} s^{-1}$, and is required to have a positive center at 850, 500 and 200hPa. There also must be a difference in the 850hPa to 200hPa vorticities of at least $6 \times 10^{-5} s^{-1}$ to provide evidence of a warm core, with a reduction in the T63 vorticity with height checked between consecutive pressure levels. These criteria must be valid for at least 1 day.

Acknowledgments. The authors would like to thank all members of the U.S. CLIVAR Hurricane Working Group for their contribution to this significant effort. We also would like to thank Naomi Henderson for making the model data available for the working group and managing the dataset. DAS, SJC and AHS acknowledge support of NSF AGS 1143959. SJC, AHS and DK acknowledge support for the GISS model runs and analysis from NASA grant NNX09AK34G. ES acknowledges support from the Italian Ministry of Education, University and Research and the Italian Ministry of Environment, Land and Sea under the GEMINA project. M.W. was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 as part of their Regional and Global Climate Modeling Program. CAM5 simulations used resources of the National Energy Research Scientific Computing Center (NERSC), also supported by the Office of Science of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. The model data used in this paper is part of the US CLIVAR Hurricane Working Group dataset.

Currently the data is only available for Working Group members, in a near future, the data will be made available for the scientific community.

Notes

1. The HadGEM3 models only tracked TCs for specific seasons (May-November for the Northern Hemisphere and October-May for the Southern Hemisphere).
2. The FSU model interannual simulation was only performed between June and November of each year and the tracking scheme was only done in the North Atlantic and North Pacific basins.
3. The GEOS-5 model used different physical parametrizations (minimum entrainment threshold for parameterized deep convection in the modified Relaxed Arakawa-Schubert convection scheme, as well as a different time step) in the climatological and interannual simulations, which led a very different TC global frequency between those runs.
4. Using the warm and cold ENSO (El Niño Southern Oscillations) definitions of the Climate Prediction Center, available at `http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml`.

References

Bell, R., J. Strachan, P. L. Vidale, K. Hodges, and M. Roberts (2013), Response of tropical cyclones to idealized climate change experiments in a global high resolution coupled general circulation model, *J. Climate*, *26*, 7966–7980, doi:10.1175/JCLI-D-12-00749.1.

Bender, M. A., T. R. Knutson, R. E. Tuleya, J. J. Sirutis, G. A. Vecchi, S. T. Garner, and I. M. Held (2010), Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes, *Science*, *327*, 454.

Bengtsson, L., H. Böttger, and M. Kanamitsu (1982), Simulation of hurricane-type vortices in a general circulation model, *Tellus*, *34*, 440–457.

Bengtsson, L., K. I. Hodges, and M. Esch (2007a), Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analysis, *Tellus*, *59 A*, 396 – 416.

Bengtsson, L., K. I. Hodges, M. Esch, N. Keenlyside, L. Kornblueh, J.-J. Luo, and T. Yamagata (2007b), How may tropical cyclones change in a warmer climate?, *Tellus*, *59 A*, 539 – 561.

Camargo, S. J. (2013), Global and regional aspects of tropical cyclone activity in the CMIP5 models, *J. Climate*, *26*, 9880–9902, doi:10.1175/JCLI-D-12-00549.1.

Camargo, S. J., and A. G. Barnston (2009), Experimental seasonal dynamical forecasts of tropical cyclone activity at IRI, *Wea. Forecasting*, *24*, 472 – 491.

Camargo, S. J., and S. E. Zebiak (2002), Improving the detection and tracking of tropical storms in atmospheric general circulation models, *Wea. Forecasting*, *17*, 1152–1162.

Camargo, S. J., A. G. Barnston, and S. E. Zebiak (2005), A statistical assessment of tropical cyclones in atmospheric general circulation models, *Tellus*, *57A*, 589–604.

Camargo, S. J., A. H. Sobel, A. G. Barnston, and P. J. Klotzbach (2010), The influence of natural climate variability on tropical cyclones and seasonal forecasts of tropical cyclone activity, in *Global Perspectives on Tropical Cyclones, from Science to Mitigation*, edited by J. C. L. Chan and J. D. Kepert, Series on Earth System Science in Asia, 2 ed., chap. 11, pp. 325–360, World Scientific.

Chavas, D. R., and K. A. Emanuel (2010), A QuickSCAT climatology of tropical cyclone size, *Geophys. Res. Lett.*, *37*, doi:10.1029/2010GL044558.

Chavas, D. R., and K. A. Emanuel (2014), Equilibrium tropical cyclone size in an idealized state of radiative-convective equilibrium, *J. Atmos. Sci.*, *71*, 1663–1680.

Chia, H. H., and C. F. Ropelewski (2002), The interannual variability in the genesis location of tropical cyclones in the Northwest Pacific, *J. Climate*, *15*, 2934–2944.

Chu, J.-H., C. R. Sampson, A. S. Levine, and E. Fukada (2002), The joint typhoon warning tropical cyclone best-tracks, 1945-2000, *Tech. Rep. NRL/MR/7540-02-16*, Naval Research Laboratory.

Dean, L., K. A. Emanuel, and D. R. Chavas (2009), On the size distribution of atlantic tropical cyclones, *Geophys. Res. Lett.*, *36*, L14,803, doi:10.1029/2009GL39051.

Emanuel, K. (2003), Tropical cyclones, *Ann. Rev. Earth Planet. Sci.*, *31*, 75–104.

Emanuel, K. A. (2013), Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century, *Proc. Nat. Acad. Sci.*, *110*, 12,219–12,224, doi: 10.1073/pnas.1301293110.

Held, I. M., and M. Zhao (2011), The response of tropical cyclone statistics to an increase in CO_2 with fixed sea surface temperatures, *J. Climate*, *20*, 5353–5364.

Hodges, K. I. (1995), Feature tracking on the unite sphere, *Mon. Wea. Rev.*, *19*, 5686–5699.

Hodges, K. I. (1996), Spherical nonparametric estimators applied to the UGAMP model integration for AMIP, *Mon. Wea. Rev.*, *124*, 2914–2932.

Hodges, K. I. (1999), Adaptive constraints for feature tracking, *Mon. Wea. Rev.*, *127*, 1362–1373.

Horn, M., K. Walsh, M. Zhao, S. J. Camargo, E. Scoccimarro, H. Murakami, H. Wang, A. Ballinger, A. Kumar, D. A. Shaevitz, J. A. Jonas, and K. Oouchi (2014), Tracking scheme dependence of simulate tropical cyclone response to idealized climate simulations, *J. Climate*, submitted.

Iizuka, S., and T. Matsuura (2008), ENSO and Western Pacific tropical cyclone activity simulated in a CGCM, *Clim. Dyn.*, *7-8*, 815–830.

JTWC (2014), Joint Typhoon Warning Center Tropical Cyclone Best Track Data Site, `http://www.npmoc.navy.mil`.

Knaff, J. A., S. P. Longmore, and D. A. Molenar (2014), An objective satellite-based tropical cyclone size climatology, *J. Climate*, *27*, 455–476.

Knutson, T. R., J. McBride, J. Chan, K. A. Emanuel, G. Holland, C. Landsea, I. Held, J. Kossin, A. K. Srivastava, and M. Sugi (2010), Tropical cyclones and climate change, *Nature Geoscience*, *3*, 157–163, doi:10.1038/ngeo779.

Landsea, C. W., and J. L. Franklin (2013), Atlantic hurricane database uncertainty and presentation of a new database format, *Mon. Wea. Rev.*, *141*, 3576–3592.

LaRow, T. E., Y.-K. Lim, D. W. Shin, E. P. Chassignet, and S. Cocke (2008), Atlantic basin seasonal hurricane simulations, *J. Climate*, *21*, 3191–3206.

Manabe, S., J. L. Holloway, and H. M. Stone (1970), Tropical circulation in a time-integration of a global model of the atmosphere, *J. Atmos. Sci.*, *27*, 580–613.

Mizuta, R., H. Yoshimura, H. Murakami, M. Matsueda, H. Endo, T. Ose, K. Kamiguchi, M. Hosaka, M. Sugi, S. Yukimoto, S. Kusunoki, , and A. Kitoh (2012), Climate simulations using the improved MRI-AGCM with 20-km grid, *J. Meteor. Soc. Japan*, *90A*, 235–260.

Murakami, H., Y. Wang, H. Yoshimura, R. Mizuta, M. Sugi, E. Shindo, Y. Adachi, S. Yukimoto, M. Hosaka, S. Kusunoki, T. Ose, , and A. Kitoh (2012), Future changes in tropical cyclone activity projected by the new high-resolution MRI-AGCM, *J. Climate*, *25*, 3237–3260.

NHC (2013), NHC (National Hurricane Center) best track dataset, available online at http://www.nhc.noaa.gov.

Prabhat, O. Rubel, S. Byna, K. S. Wu, M. Wehner, and W. Bethel (2012), TECA: A parallel toolkit for extreme climate analysis, in *Proc. Int. Conf. on Computational Science,* *ICCS 2012*, *Procedia Computer Science*, vol. 9, edited by H. Ali, Y. Shi, D. Khazanchi, M. Lees, G. VanAlbada, J. Dongarra, and P. M. A. Sloot, pp. 866–876.

Rayner, N., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 2156–2202.

Rienecker, M. M., M. J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R. D. Koster, R. Gelaro, I. Stajner, , and J. E. Nielsen (2008), The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0, *Technical Report Series on Global Modeling and Data Assimilation, Volume 27 TM-2008-104606*, NASA.

Roeckner, E., G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, L. Kornblueh, U. Schlese, U. Schulzweida, I. Kirchner, E. Manzini, A. Rhodin, and A. Tompkins (2003), The atmospheric general circulation model ECHAM5. Part I: Model description, *Tech. Rep. 349*, Max-Planck Institute for Meteorology, Hamburg, Germany, 127 pp.

Saffir, H. S. (1977), Design and construction requirements for hurricane resistant construction, *Tech. Rep. Preprint No. 2830, 20 pp.*, ASCE, available from American Society of Civil Engineers, New York, NY 10017.

Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y.-T. Hou, H.-Y. Chuang, M. Iredell, M. Ek, J. Meng, R. Yang, M. P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E. Becker (2014), The NCEP climate forecast system version 2, *J. Climate*, *27*, 2185–2208.

Schmidt, G. A., M. Kelley, L. Nazarenko, R. Ruedy, G. L. Russell, I. Aleinov, M. Bauer, S. E. Bauer, M. K. Bhat, R. Bleck, V. Canuto, Y. H. Chen, Y. Cheng, T. L. Clune, A. D. Genio, R. de Fainchtein, G. Falugevi, J. E. Hansen, R. J. Healy, N. Y. Kiang, D. Koch, A. A. Lacis, A. N. LeGrande, J. Lerner, K. K. Lo, E. E. Matthews, S. Menon, R. L. Miller, V. Oinas, A. O. Oloso, J. P. Perlwitz, M. J. Puma, W. M. Putman, D. Rind, A. Romanou, M. Sato, D. T. Schindell, S. Sun, R. A. Syed, N. Tausnev, K. Tsigaridis, N. Unger, A. Voulgarakis, M.-S. Yao, and J. Zhang (2014), Configuration and assessment of GISS ModelE2 contributions to the CMIP5 archive, *J. Adv. Model. Earth Sys.*, *6*, 141–184.

Scoccimarro, E., S. Gualdi, A. Bellucci, A. Sanna, P. G. Fogli, E. Manzini, M. Vichi, P. Oddo, and A. Navarra (2011), Effects of tropical cyclones on ocean heat transport in a high resolution coupled general circulation model, *J. Clim.*, *24*, 4368–4384.

Scoccimarro, E., S. Gualdi, G. Villarini, G. A. Vecchi, M. Zhao, K. Walsh, and A. Navarra (2014), Increased precipitation events associated with landfalling tropical cyclones in response to a warmer climate and increased co_2, *J. Climate*, *early online*, doi:10.1175/JCLI-D-14-00065-1.

Strachan, J., P. L. Vidale, K. Hodges, M. Roberts, and M.-E. Demory (2013), Investigating global tropical cyclone activity with a hierarchy of AGCMs: the role of model resolution, *J. Climate*, *26*, 133–152.

Strazzo, S., J. B. Elsner, T. LaRow, and D. J. Halperin (2013), Observed versus GCM-generated local tropical cyclone frequency: Comparisons using a spatial lattice, *J. Climate*, *26*, 8257–8268.

Villarini, G., D. A. Lavers, E. Scoccimarro, M. Zhao, M. F. Wehner, G. A. Vecchi, T. R. Knutson, and K. A. Reed (2014), Sensitivity of tropical cyclone rainfall to idealized global scale forcings, *J. Climate*, doi:10.1175/JCLI-D-13-00780.1, early online.

Vitart, F. (1998), Tropical storm interannual and interdecadal variability in an ensemble of GCM integrations, Ph.D. thesis, Princeton University, 387 pp.

Vitart, F., J. L. Anderson, and W. F. Stern (1997), Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations, *J. Climate*, *10*, 745–760.

Vitart, F., D. Anderson, and T. Stockdale (2003), Seasonal forecasting of tropical cyclone landfall over Mozambique, *J. Climate*, *16*, 3932–3945.

Vitart, F., M. R. Huddleston, M. Déqué, D. Peake, T. N. Palmer, T. N. Stockdale, M. K. Davey, S. Inenson, and A. Weisheimer (2007), Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in june byEUROSIP, *Geophys. Res. Lett.*, *34*, L16,815, doi:10.1029/2007GL030740.

Walsh, K. (1997), Objective detection of tropical cyclones in high-resolution analyses, *Mon. Wea. Rev.*, *125*, 1767–1779.

Walters, D., M. Best, A. Bushell, D. Copsey, J. Edwards, P. Falloon, C. Harris, A. Lock, J. Manners, C. Morcrette, M. Roberts, R. Stratton, S. Webster, J. Wilkinson, M. Willett, I. Boutle, P. Earnshaw, P. Hill, C. MacLachlan, G. Martin, W. Moufouma-Okia, M. Palmer, J. Petch, G. Rooney, A. Scaife, and K.D.Williams (2011), The Met Office Unified Model Global Atmosphere 3.0/3.1 and JULES Global Land 3.0/3.1 configurations, *Geosci. Model Dev.*, *4*, 919–941, doi:10.5194/gmd-4-919-2011.

Wehner, M. F., K. Reed, F. Li, Prabhat, J. Bacmeister, C.-T. Chen, C. Paciorek, P. Gleckler, K. Sperber, W. D. Collins, A. Gettelman, C. Jablonowski, and C. Algieri (2014), The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1, *J. Adv. Model. Earth Sys.*, *submitted*.

Zhao, M., and I. M. Held (2010), An analysis of the effect of global warming on the intensity of atlantic hurricanes using a GCM with statistical refinement, *J. Climate*, *23*, 6382–6393.

Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi (2009), Simulations of global hurricane climatology, interannual variability and response to global warming using a 50 km resolution gcm, *J. Climate*, *22*, 6653 – 6678.

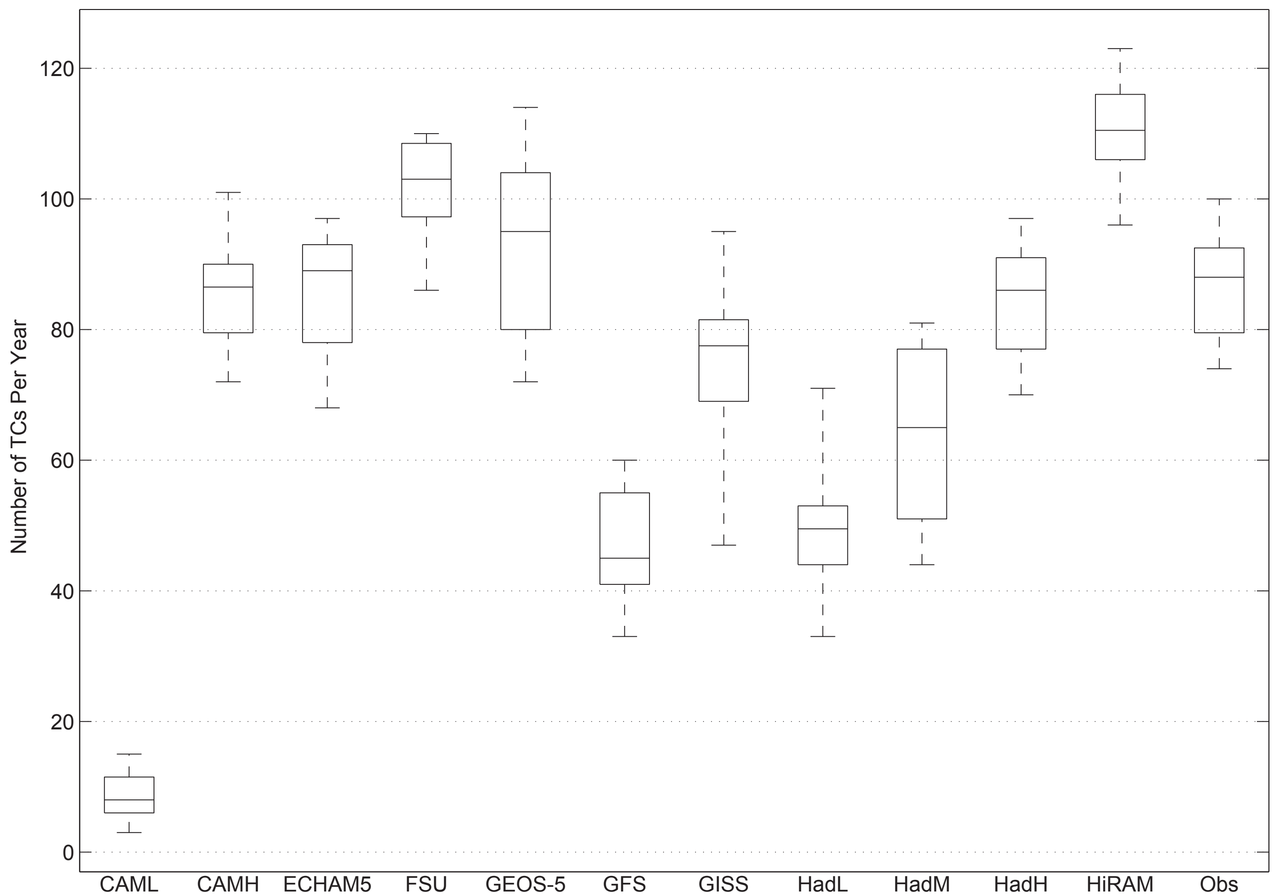

Figure 1. Distributions of the number of TCs per year for models and observations. The horizontal line inside the boxes shows the median number of TCs per year, the top and bottom of the boxes represent the 75th and 25th percentiles respectively, with the whiskers extending to the maximum and minimum number of TCs per year in each case. CAML: Low-resolution CAM5.1, CAMH: High-resolution CAM5.1, HadL: Low-resolution HadGEM3, HadM: Medium-resolution HadGEM3, HadH: High-resolution HadGEM3.

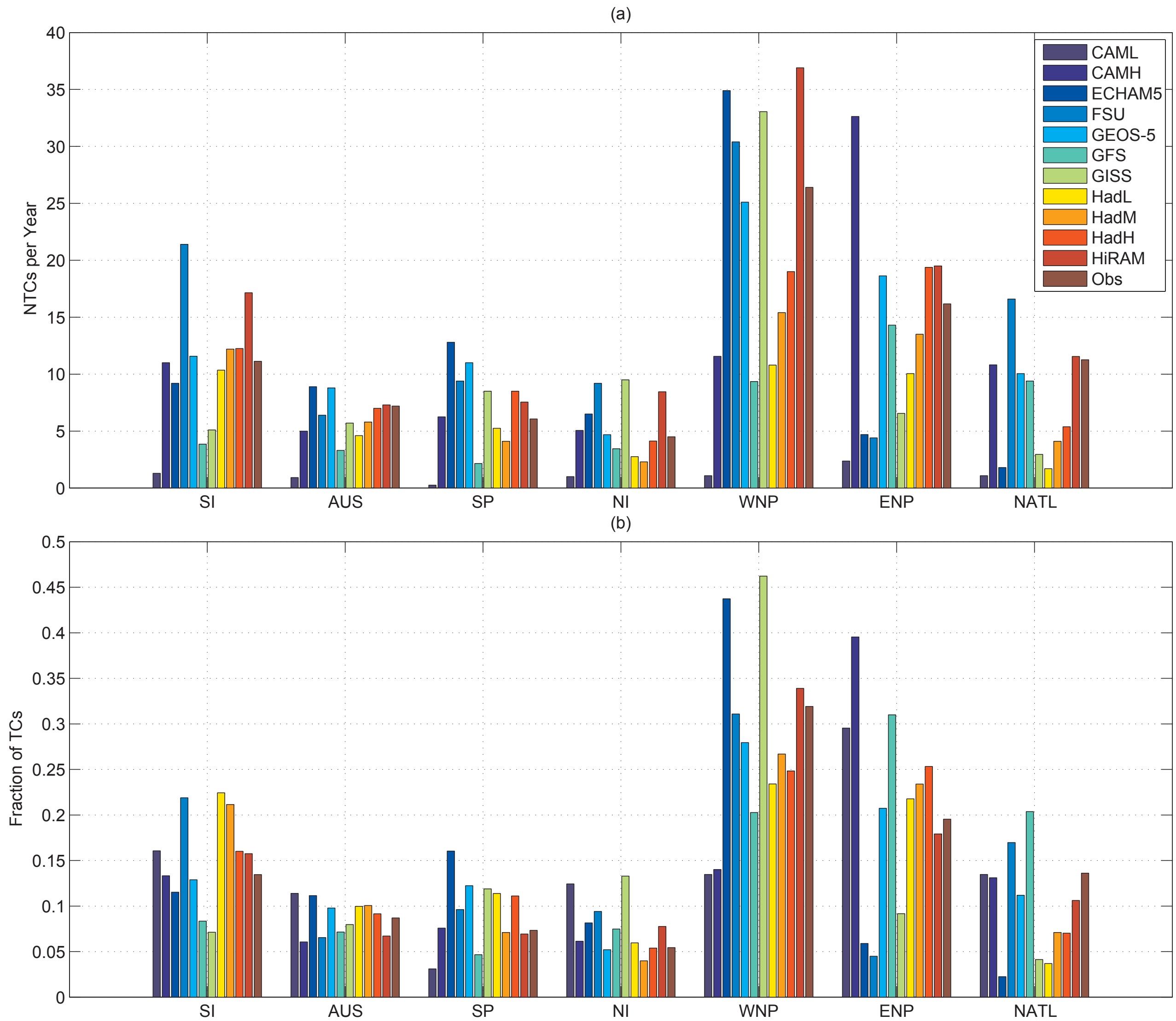

Figure 2. Mean number of TCs formed in each basin for models and observations. (a) shows the total number of TCs, (b) shows the percentage of TCs in each basin. The basins are defined as: SI (South Indian), AUS (Australian), SP (South Pacific), NI (North Indian), WNP (Western North Pacific), ENP (Easter North Pacific), NATL (North Atlantic). The model names follow the definitions in Fig. 1.

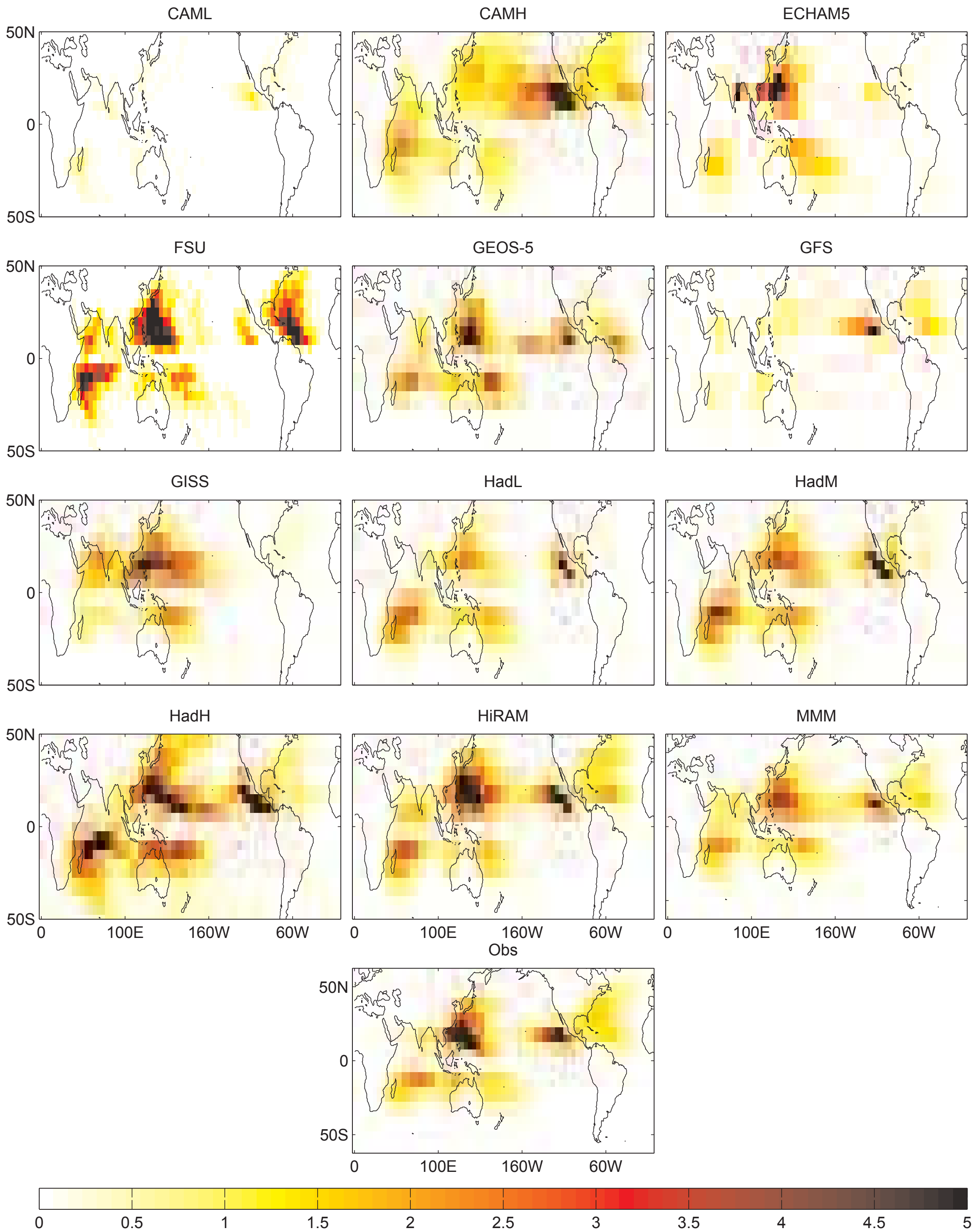

Figure 3. TC track density in models and observations. Track density is defined as the number of TC transits per 5° x 5° box per year. The total number of transits in each grid point and model is obtained and then divided by the number of years in each model simulation. The multi-model mean (MMM) track density is also shown. In the case of CAM5.1 and HadGEM3 only the high-resolution version was included in the multi-model mean.

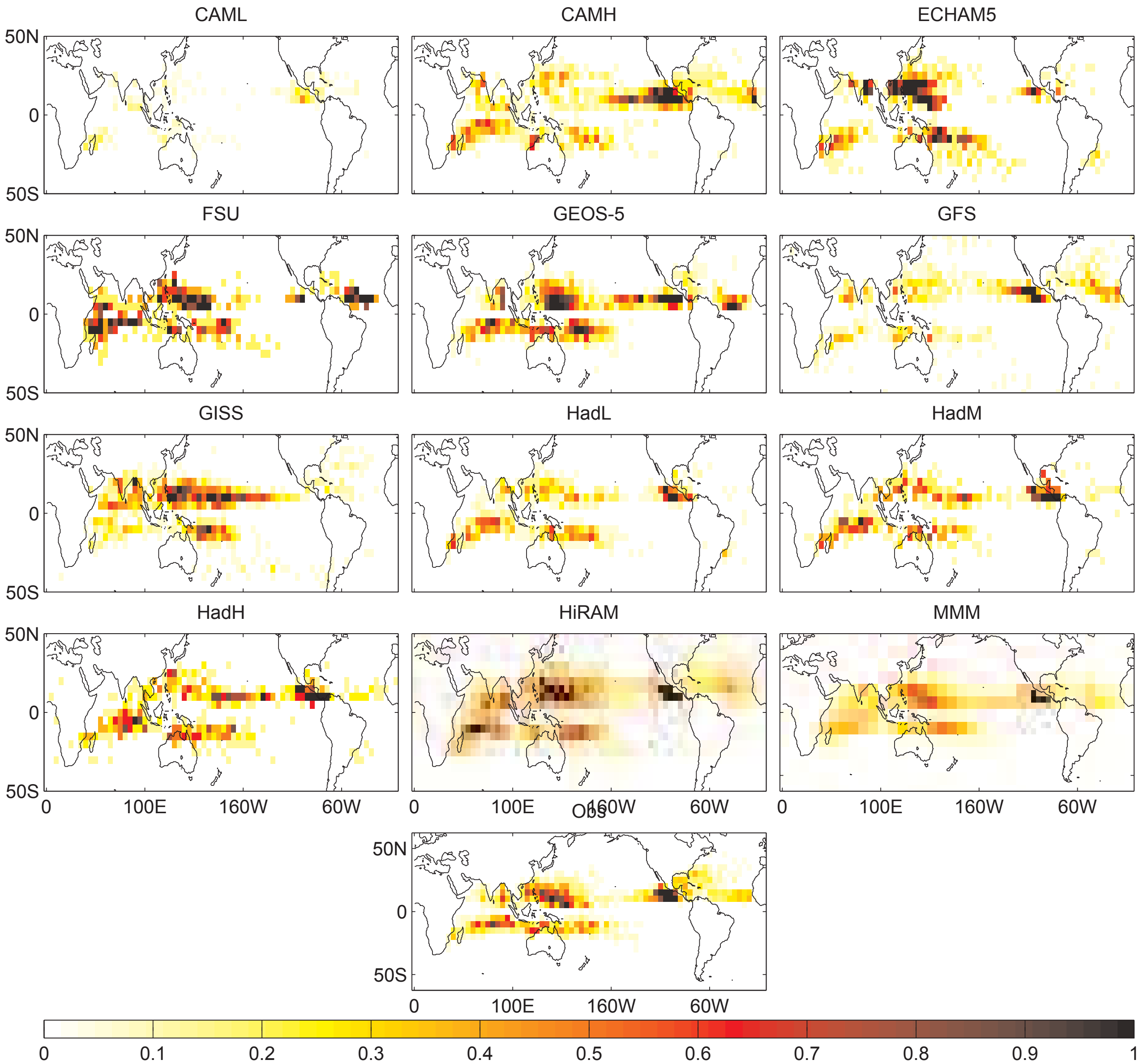

Figure 4. TC genesis density in models and observations. Genesis density is defined as the number of TC formation per 5° x 5° box per year. The total number of transits in each grid point and model is obtained and then divided by the number of years in each model simulation. The multi-model mean (MMM) track density is also shown. In the case of CAM5.1 and HadGEM3 only the high-resolution version was included in the multi-model mean.

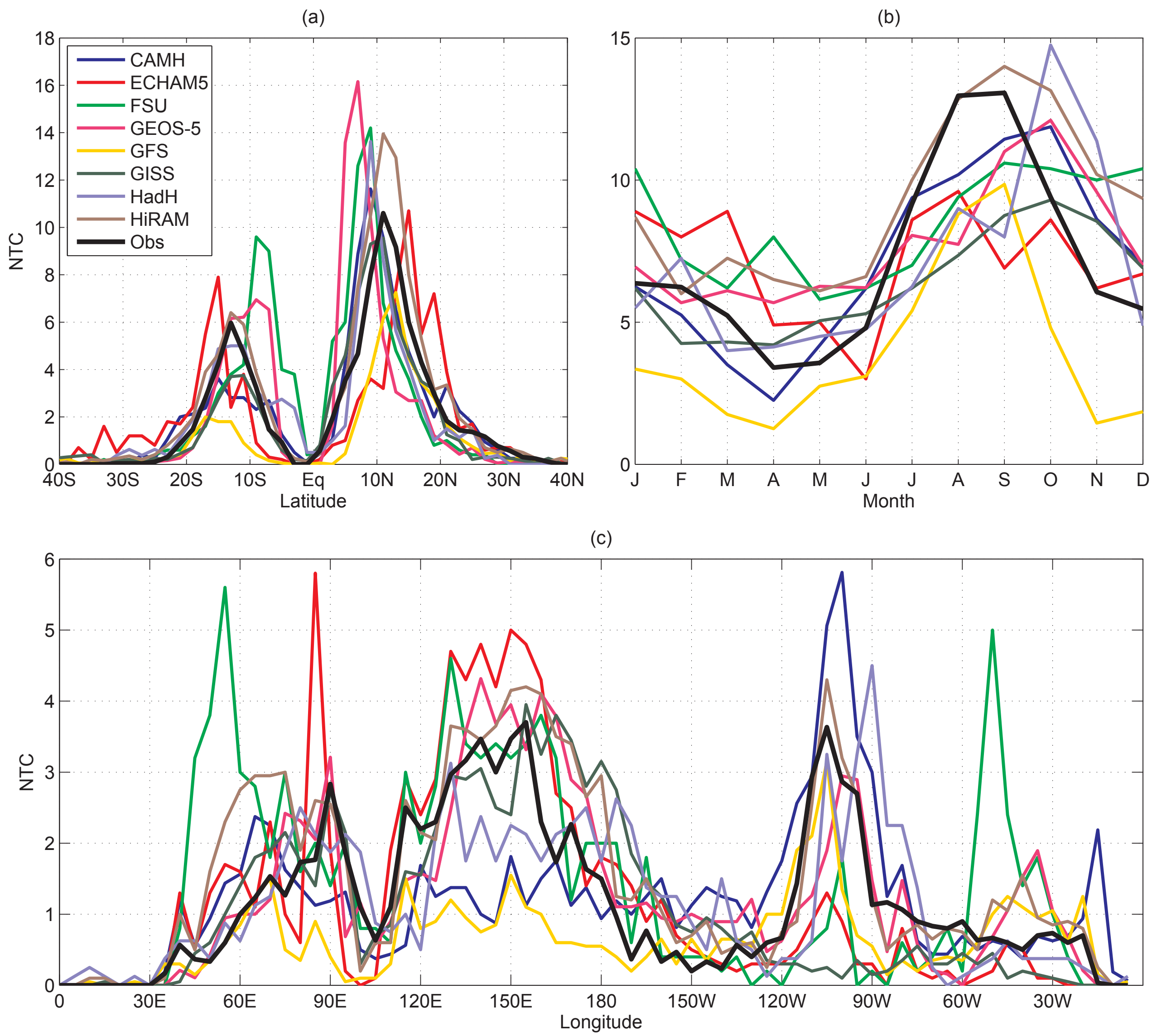

Figure 5. Mean number of TC genesis per year in models and observations as a function of latitude (a), month (b), and longitude (c). The latitude (longitude) counts are per 2 (5) degrees bins.

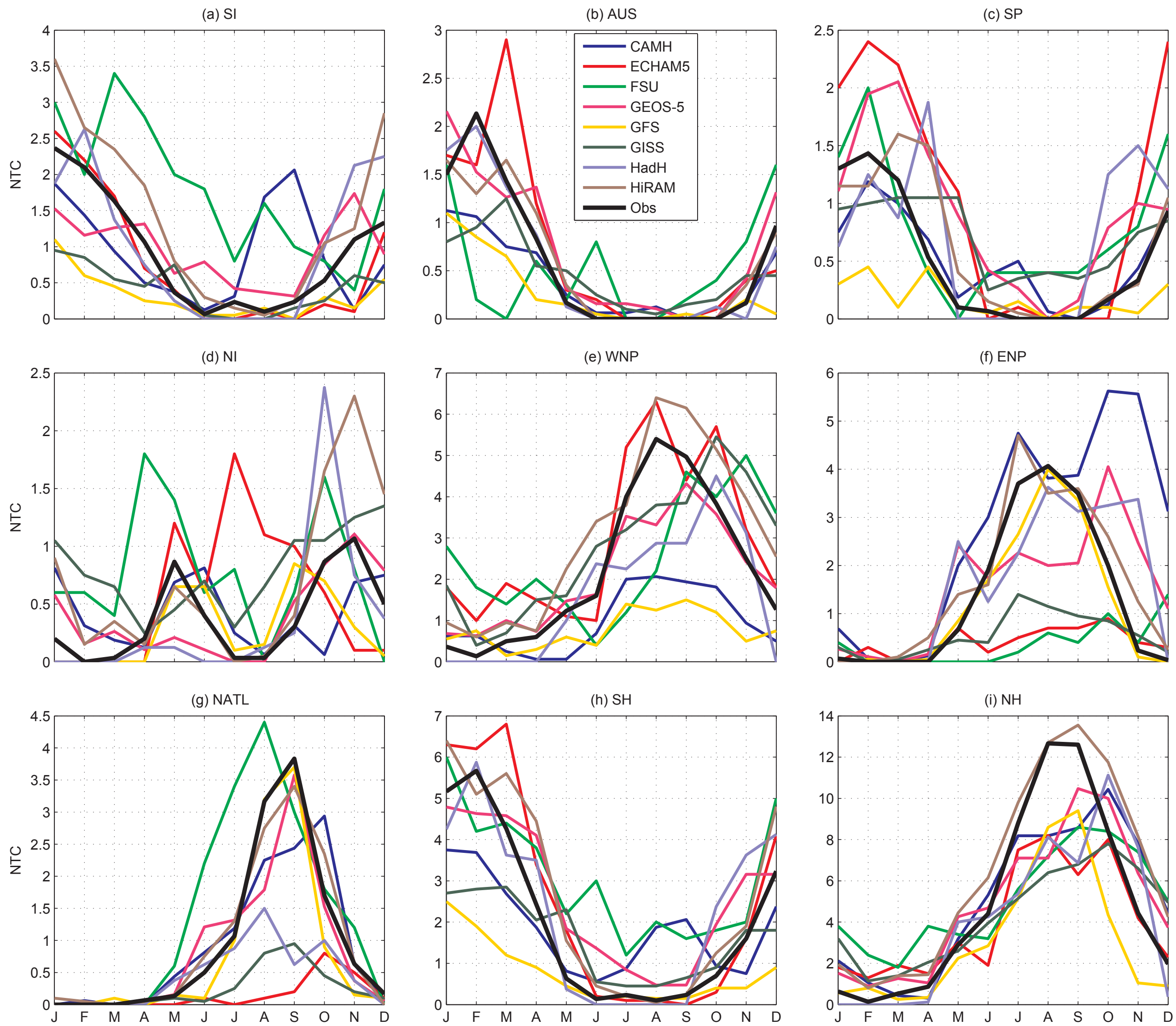

Figure 6. Mean TC genesis per year and month in models and observation in various basins (as defined in Fig. 2) and in the southern (SH) and northern (NH) hemispheres

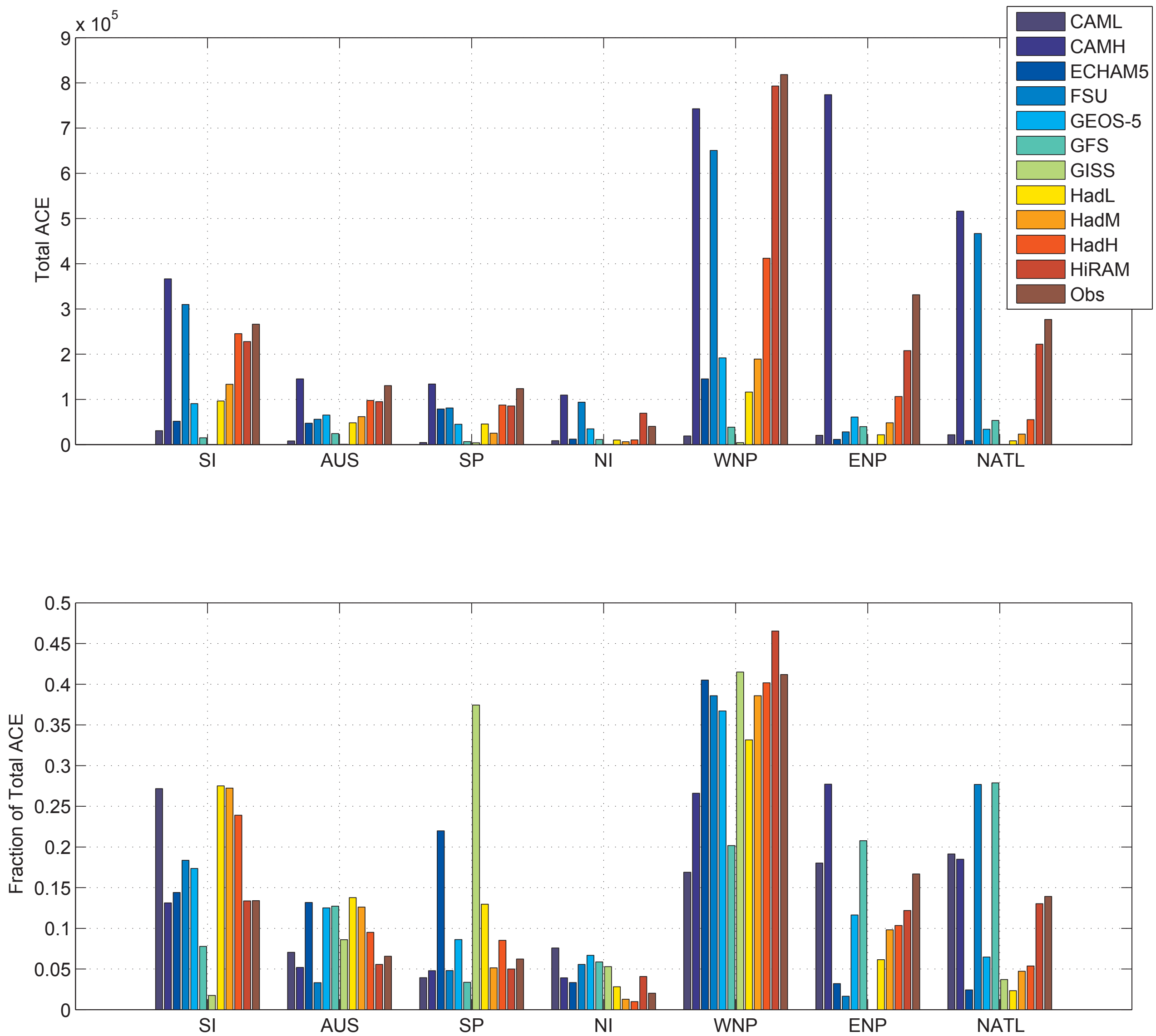

Figure 7. Accumulated cyclone energy (ACE) for models and observations (top panel). The bottom panel shows the percentage of the ACE in each basin for models and observations. Basins and models are defined as in previous figures.

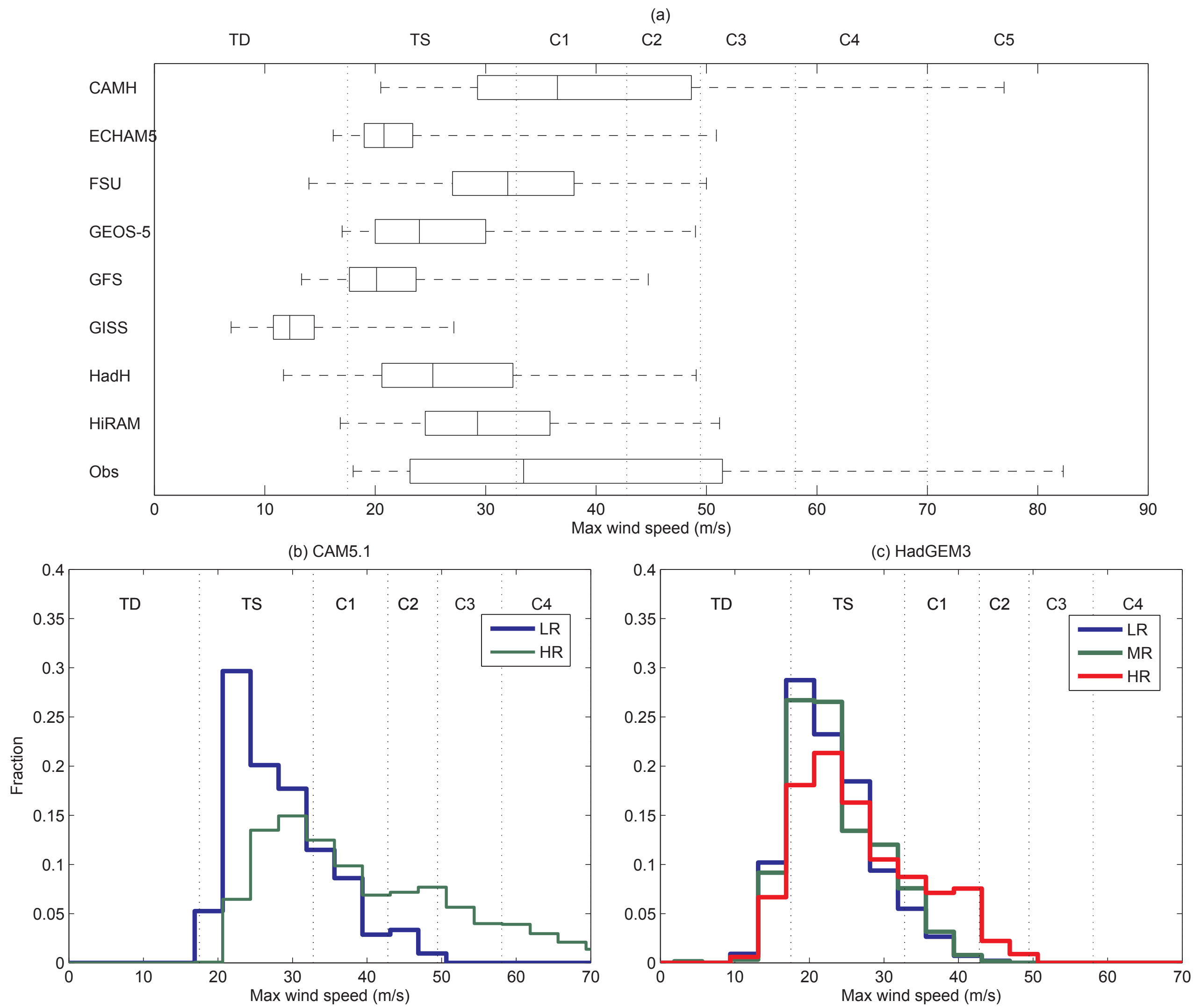

Figure 8. Distributions of TC maximum intensity in models and observations (a). The vertical line shows the median of each distribution, the left and right edges of the box represent the 75th and 25th percentiles respectively, and the whiskers extend to the maximum and minimum values in each case. Histograms of TC maximum intensity for two horizontal resolutions of the CAM5.1 model (b) and three model resolutions of the HadGEM1 model (c). The vertical lines show the boundaries of the Saffir-Simpson hurricane classification scale. TD: Tropical Depression, TS: Tropical Storm, C1-C5: Category 1-5 hurricanes. LR: Low resolution, MR: Medium resolution, HR: High resolution.

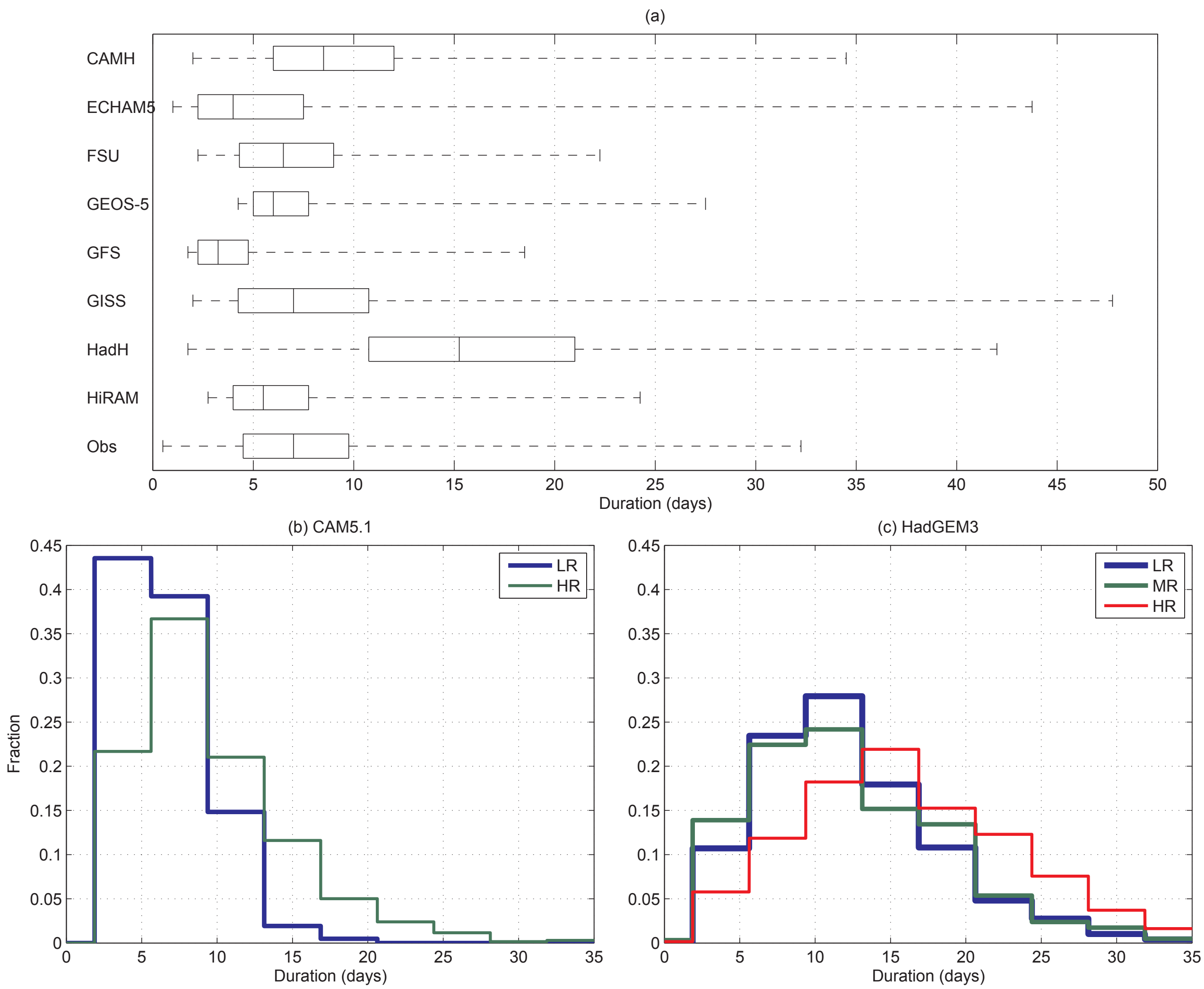

Figure 9. (a) Distributions of TC lifetime (or duration) for models and observations. The vertical line shows the median of each distribution, the left and right edges of the box represent the 75th and 25th percentiles respectively, and the whiskers extend to the maximum and minimum values in each case. The histograms of TC durations in the CAM5.1 and HadGEM3 models for different resolutions are shown in (b) and (c), respectively. LR: Low resolution, MR: Medium resolution, HR: High resolution.

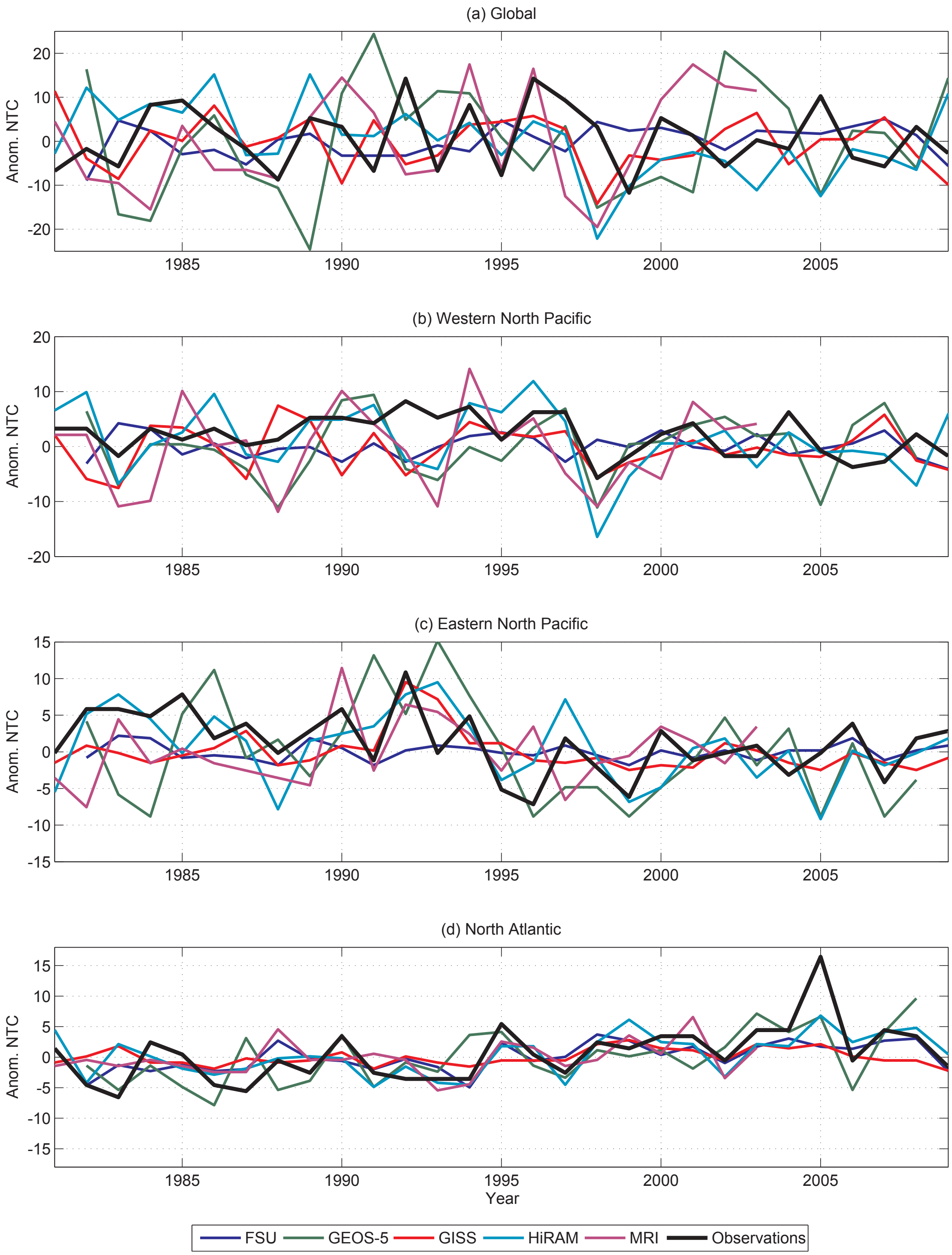

Figure 10. Anomalous number of TCs per year (number of TCs per year minus the mean number of TCs for all years) (a) in the globe and in a few of the Northern Hemisphere basins (Western North Pacific (b), Eastern North Pacific (c), and North Atlantic (d)). For the models, when more than one ensemble simulation was available, the ensemble mean anomlay number of TCs in each year is shown.

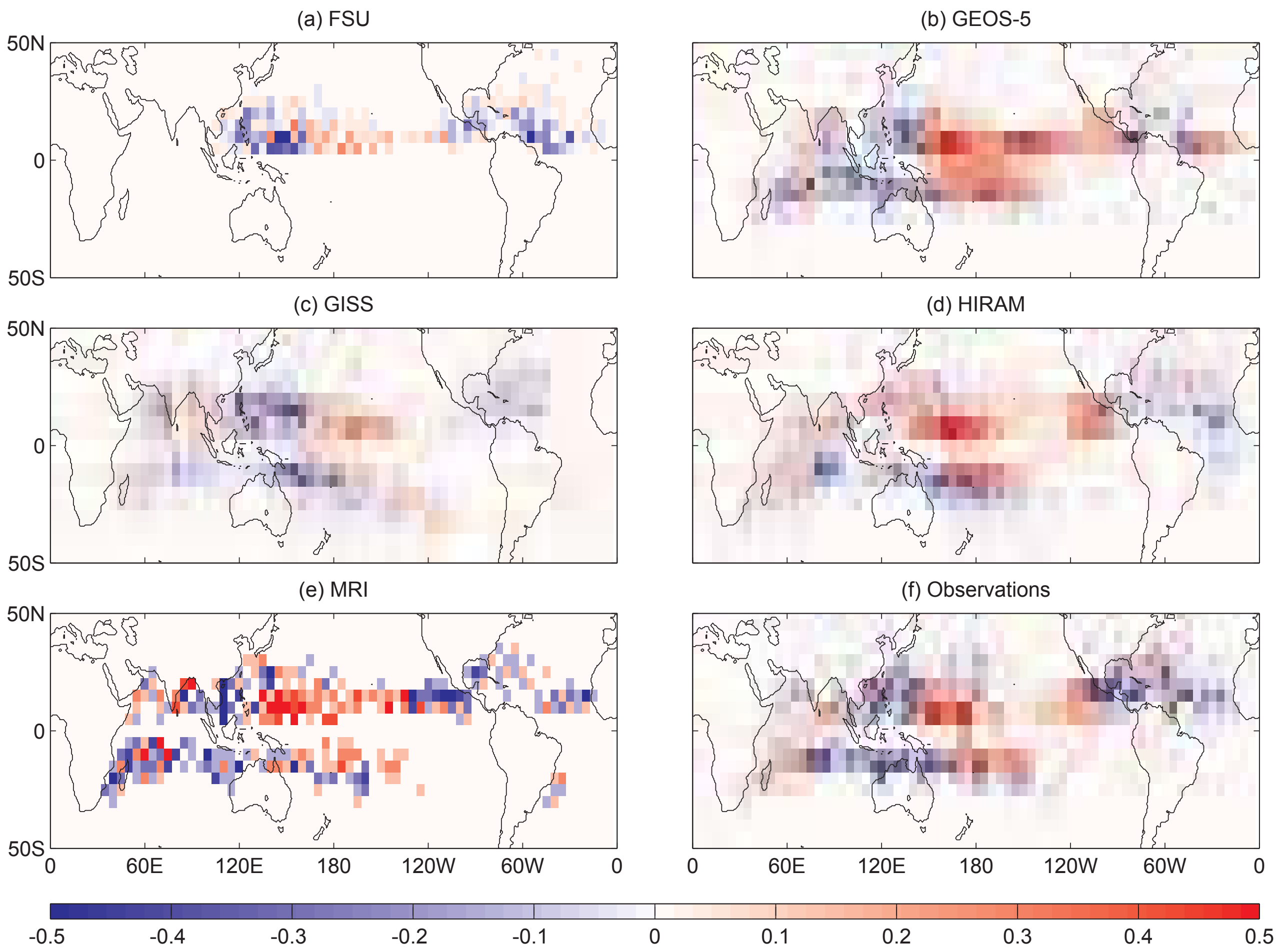

Figure 11. Difference of TC genesis density in El Niño and La Niña in models and observations. The genesis density is defined as the mean TC formation per 5° x 5° box per year.

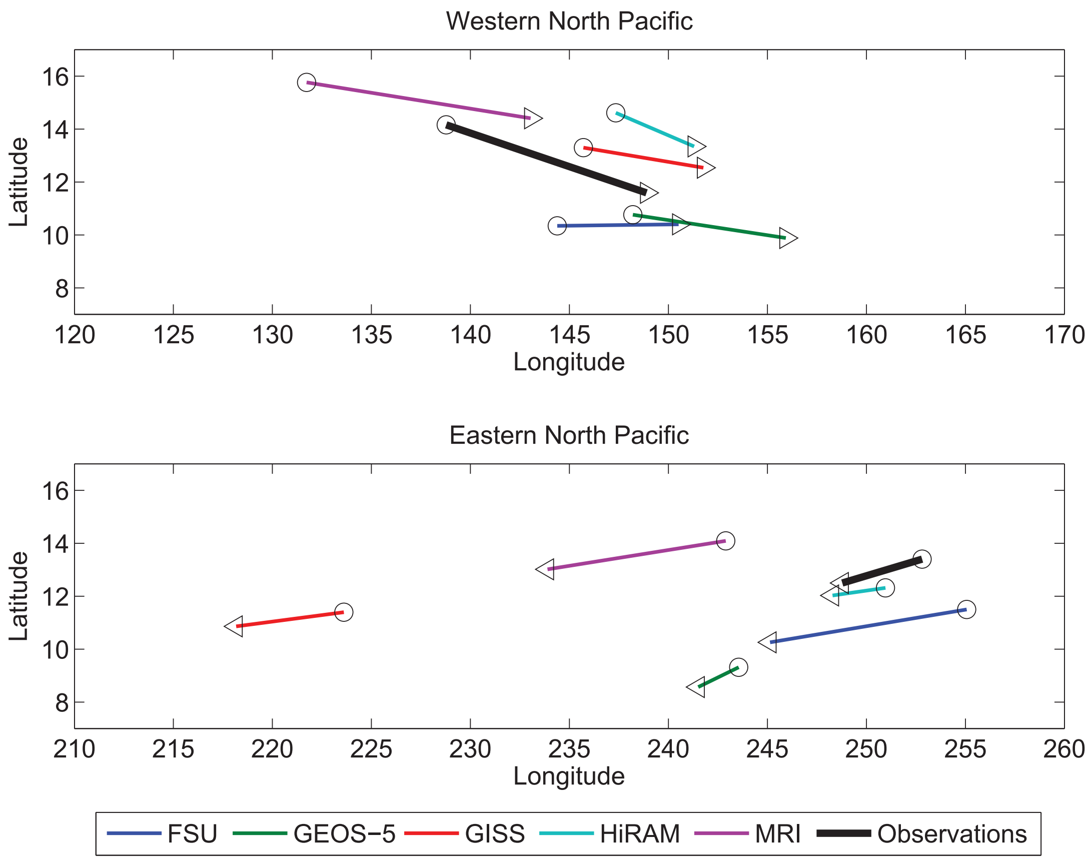

Figure 12. Mean TC genesis location in the western and eastern North Pacific in El Niño (triangles) and La Niña (circles) years in models and observations.

Table 1. Models characteristics and references for models and tracking schemes. LR: Low Resolution, MR: Medium Resolution, HR: High Resolution. References: Wehner: *Wehner et al.* [2014]; Prabhat: *Prabhat et al.* [2012]; RS: *Roeckner et al.* [2003] and *Scoccimarro et al.* [2011]; Walsh: *Walsh* [1997]; LaRow: *LaRow et al.* [2008]; Vitart: *Vitart et al.* [2003]; Rienecker: *Rienecker et al.* [2008]; Saha: *Saha et al.* [2014]; Zhao: *Zhao et al.* [2009]; Schmidt: *Schmidt et al.* [2014]; Camargo: *Camargo and Zebiak* [2002]; Walters: *Walters et al.* [2011]; HB: *Hodges* [1995] and *Bengtsson et al.* [2007a, b]; MM: *Mizuta et al.* [2012] and *Murakami et al.* [2012]; Murakami: *Murakami et al.* [2012].

Model	Resolution	Approx Res (km)	Reference	Tracking Scheme
LR CAM5.1	100 km	100	Wehner	Vitart/Prabhat
HR CAM5.1	1/4°	28	Wehner	Vitart/ Prabhat
ECHAM5	T159	84	RS	Vitart/Walsh
FSU	T126	106	LaRow	Vitart/Zhao
GEOS-5	1/2°	56	Rienecker	Vitart/Zhao
GFS	T126	106	Saha	Vitart/Zhao
GISS	1°	111	Schmidt	Camargo
LR HadGEM3	N96	130	Walters	HB
MR HadGEM3	N216	60	Walters	HB
HR HadGEM3	N320	40	Walters	HB
HiRAM	50 km	50	Zhao	Vitart/Zhao
MRI	TL319	60	MM	Murakami

Table 2. Number of years in the climatological simulations for each model.

Model	Years
LR CAM5.1	24
HR CAM5.1	16
ECHAM5	9
FSU	5
GEOS-5	20
GFS	20
GISS	20
LR HadGEM3	20
MR HadGEM3	10
HR HadGEM3	9
HiRAM	20

Table 3. Models' interannual simulations ensemble members and years.

Model	Number of Ensembles	Years
FSU	3	1982-2009
GEOS-5	2	1982-2009
GISS	3	1981-2009
HiRAM	3	1981-2009
MRI	1	1981-2003

Table 4. Correlations of yearly ACE and model-ACE in each basin (r_{AA}) and correlations the yearly observed ACE the model modified MACE (r_{AM}). The correlations are shown as r_{AA}/r_{AM}. Asterisks denote correlations that are statistically significant. Basins are defined as: SI (South Indian), AUS (Australian), SP (South Pacific), NI (North Indian), WNP (Western North Pacific), ENP (Easter North Pacific), NATL (North Atlantic).

Model	SI	AUS	SP	NI	WNP	ENP	NATL
FSU	-	-	-	-	0/0	0.5*/0.5*	0.7*/0.7*
GEOS-5	0/0	-0.1/-0.2	0.5*/0.4*	-0.2/-0.2	0.7*/0.7*	0.4*/0.5*	0.6*/0.6*
GISS	0/0	-0.3/0	-0.2/-0.2	-0.2/0.2	0.3/0.2	0/0.7*	0/0.4
HiRAM	0.2/0.2	0.4*/0.4*	0.1/0.1	-0.1/-0.1	0.5*/0.5*	0.6*/0.6*	0.7*/0.7*
MRI	0.2/0.2	-0.4*/-0.4*	0.1/0.1	-0.1/-0.1	0.3/0.3	0.4*/0.4*	0.6*/0.6*

Table 5. El Niño and La Niña seasons for the northern and southern hemispheres, using the warm and cold ENSO (El Niño Southern Oscillations) definitions of Climate Prediction Center. The northern (southern) hemisphere seasons definitions as based on the state of ENSO in the August - October (January - March) seasons. Note that the southern hemisphere TC seasons are defined from July to June, emcompassing 2 calendar years.

Northern Hemisphere		Southern Hemisphere	
El Niño	La Niña	El Niño	La Niña
1982	1983	1982/83	1980/81
1986	1985	1986/87	1984/85
1987	1988	1987/88	1988/89
1991	1995	1991/92	1995/96
1994	1998	1994/95	1998/99
1997	1999	1997/98	1999/00
2002	2000	2002/03	2000/01
2004	2007		2005/06
2006			2007/08
2009			

Table 6. Correlations of NTC per year or season (southern hemispere) in the globe by basins. Basins are defined as: SI (South Indian), AUS (Australian), SP (South Pacific), NI (North Indian), WNP (Western North Pacific), ENP (Easter North Pacific), NATL (North Atlantic). Asterisks denote correlations that are statistically significant.

Model	Global	SI	AUS	SP	NI	WNP	ENP	NATL
FSU	-0.13	–	–	–	0	-0.25	0.42	0.61*
GEOS-5	-0.21	0.20	0.07	0.32	-0.10	0.24	0.27	0.61*
GISS	-0.01	0.12	0.15	-0.26	-0.12	0.21	0.42	0.45
HiRAM	0.22	0.34	0.39	0.07	0.11	0.55*	0.51*	0.69*
MRI	0.15	0.36	0.32	-0.02	-0.37	0.35	0.22	0.55*