Change in the Odds of Warm Years and Seasons Due to Anthropogenic Influence on the Climate

NIKOLAOS CHRISTIDIS AND PETER A. STOTT
Met Office Hadley Centre, Exeter, United Kingdom

(Manuscript received 17 September 2013, in final form 13 December 2013)

ABSTRACT

The new Hadley Centre system for attribution of weather and climate extremes provides assessments of how human influence on the climate may lead to a change in the frequency of such events. Two different types of ensembles of simulations are generated with an atmospheric model to represent the actual climate and what the climate would have been in the absence of human influence. Estimates of the event frequency with and without the anthropogenic effect are then obtained. Three experiments conducted so far with the new system are analyzed in this study to examine how anthropogenic forcings change the odds of warm years, summers, or winters in a number of regions where the model reliably reproduces the frequency of warm events. In all cases warm events become more likely because of human influence, but estimates of the likelihood may vary considerably from year to year depending on the ocean temperature. While simulations of the actual climate use prescribed observational data of sea surface temperature and sea ice, simulations of the nonanthropogenic world also rely on coupled atmosphere–ocean models to provide boundary conditions, and this is found to introduce a major uncertainty in attribution assessments. Improved boundary conditions constructed with observational data are introduced in order to minimize this uncertainty. In more than half of the 10 cases considered here anthropogenic influence results in warm events being 3 times more likely and extreme events 5 times more likely during September 2011–August 2012, as an experiment with the new boundary conditions indicates.

1. Introduction

Research on detection and attribution of climate change has provided a wealth of evidence for significant anthropogenic warming in recent decades on global, continental, and subcontinental scales (Jones et al. 2013; Lewis and Karoly 2013; Gillett et al. 2012; Stott et al. 2010; Hegerl et al. 2007; Stott 2003). In a warming climate, temperature extremes are also expected to intensify both in frequency and intensity and significant human influence on changes of such characteristics has already been identified (Wen et al. 2013; Morak et al. 2013; Seneviratne et al. 2012; Christidis et al. 2011). Pushing the scientific boundaries further, new research focuses not only on long-term changes in extremes, but also on the attribution of specific high-impact events (Stott et al. 2013). In the aftermath of catastrophic events, possible links between their occurrence and external climatic forcings such as greenhouse gas emissions are invariably questioned and attribution studies aim to provide reliable and scientifically robust answers sought by both the public and policy makers.

While a specific weather and climate extreme event cannot be solely attributed to a single cause, it is still possible to estimate how certain factors like the effect of anthropogenic forcings may have modified the odds of the event. For example, Stott et al. (2004) found that human influences on the climate have more than doubled the chances of a European heat wave as severe as the one in 2003. More recently, a new ensemble-based methodology for attribution of climate-related events (ACE) has gained popularity, originally introduced by Pall et al. (2011) in a study of the floods in the United Kingdom (UK) in autumn 2000. Two large ensembles of simulations are produced with a global atmospheric model, one with and one without the effect of a climate change driver, which in this paper is the overall effect of anthropogenic forcings. The odds of an event (usually defined based on a threshold exceedance; for example, the exceedance of a climatologically high temperature in
the case of warm extremes) can then be estimated in the “actual” and the hypothetical “natural” world from the two ensembles and so the change in the odds due to anthropogenic forcings can be inferred. A new attribution system that facilitates this ACE methodology was developed in the Hadley Centre (Christidis et al. 2013b) to assess the impact of human influence on extremes.

The ACE methodology discussed above requires prescribed values of sea surface temperature (SST) and sea ice as boundary conditions in the model experiments. In simulations of the actual climate these come from observational data. In the case of the natural climate, estimates of the anthropogenic change in the SST and sea ice are subtracted from the observations, which typically come from experiments with an atmosphere–ocean coupled general circulation model (GCM). The representation of the SST in the natural world is a major uncertainty in the methodology and is often accounted for by generating several ensembles (i.e., different versions) of the natural world, using estimates of the change in the SST from different GCMs. On the other hand, the use of computationally cheaper atmospheric models enables the generation of large ensembles, an important advantage when computing the statistics of rare events. Moreover, by prescribing the SST, we retain modes of variability (such as phases of the El Niño–Southern Oscillation) that may in some cases be key drivers of regional extremes (e.g., La Niña years associated with increased risk of extreme winter precipitation in the greater horn of Africa) (Arribas et al. 2011). If the occurrence of regional extremes is largely influenced by the SST, the odds of having an event may vary considerably from year to year due to the observed state of the SSTs, much of which will be affected by natural internal variability, and an important question is how anthropogenic forcings may change the odds of an event given a certain oceanic state. This is different from the European heat wave study of Stott et al. (2004), which provided the odds of a heat wave event in the general case where any SST patterns may occur.

In this paper we take advantage of the three experiments that have so far been carried out with the new Hadley Centre ACE system (Christidis et al. 2013b), which cover three recent consecutive years. The aim of our study is twofold. First, we estimate how the odds of warm years and seasons have changed due to human influence on the climate given the observed patterns of SST associated with internal variability and other factors. We make these calculations for the period covered by the three experiments (September 2010–August 2012) in a number of subcontinental regions. Such assessments are a key output of the ACE system and need to be made on an annual (or a case by case) basis, as the oceanic conditions in the reference regions constantly change. Second, we examine the sensitivity of the attribution assessments on the SST patterns used in simulations of the natural world and introduce improved patterns derived from observational data rather than GCMs. In the remainder of the paper we discuss the attribution system and data used for this study (section 2), present attribution assessments for different regions (section 3), introduce the new SST representation in the simulations without anthropogenic forcings (section 4), and briefly discuss the main findings and implications of this work (section 5).

2. Attribution experiments with the Hadley Centre ACE system

Our ACE system is built on the Hadley Centre Global Environmental Model, version 3–Atmospheric (HadGEM3-A), currently run at a horizontal resolution of 1.25° longitude by 1.875° latitude with 38 vertical levels (Hewitt et al. 2011). We generate ensembles by introducing random perturbations in physical parameters (Murphy et al. 2004) as well as wind increments to account for grid-scale sources of kinetic energy (Tennant et al. 2011). Simulations of the actual climate include all the main anthropogenic and natural forcings detailed in Stott et al. (2006). Anthropogenic forcings comprise historical changes in well-mixed greenhouse gases, tropospheric and stratospheric ozone, sulfate aerosols, black carbon, aerosols from biomass burning, and changes in land use. Simulations without the effect of human influence include only natural forcings (i.e., changes in volcanic aerosols and the solar irradiance). The attribution system is described in detail in Christidis et al. (2013b) and has already been used to study numerous high impact extremes, including the catastrophic heat wave in Moscow in July 2010, the two consecutive cold winters of 2009/10 and 2010/11 in the UK, the devastating East African drought in 2011, and the severe floods in eastern Australia in March 2012 (Christidis et al. 2013a,b; Christidis and Stott 2012; Lott et al. 2013).

Three experiments have so far been carried out since the setup of the attribution system to enable investigations of extreme events in three consecutive years. Details are summarized in Table 1. Each experiment comprises an ensemble of simulations with all forcings (ALL) that provides different realizations of the actual world, as well as ensembles without the effect of human influence (NAT) for the natural world. The ALL simulations employ SST and sea ice data from the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST) dataset (Rayner et al. 2003). The NAT simulations use a GCM estimate of the change in the SST (ΔSST) due to anthropogenic forcings, which is subtracted.
from the HadISST data. ΔSST is calculated from the difference between GCM runs with all forcings and runs with natural forcings only, or, alternatively, the difference between runs with anthropogenic forcings and the mean of a long control GCM run without any external forcings on the climate (Christidis et al. 2013b). Estimates of ΔSST are produced for each month separately. For each experiment we have three different NAT ensembles, generated with ΔSST estimates from different GCMs to help assess the sensitivity of the results to the SST representation in simulations of the natural climate. Each experiment spans a full year from September to August, rather than January–December, and for that reason we calculate annual means in our study as the September–August mean. The three years covered by the ACE experiments (Table 1) are henceforth referred to as years 1, 2, and 3.

While all GCMs used to estimate ΔSST indicate an overall anthropogenic warming of the oceans, there are distinct differences both in the regional characteristics and the magnitude of the change. This is illustrated in Fig. 1, which shows the annual mean anthropogenic SST change used in the experiment for year 3 computed with three different models. The Hadley Centre Global Environment Model, version 2–Earth System (HadGEM2-ES; Jones et al. 2011) Hadley Centre Global Environment Model, version 1 (HadGEM1; Stott et al. 2006) Hadley Centre Coupled Model, version 3 (HadCM3; Johns et al. 2003) HadGEM2-ES (Jones et al. 2011) HadGEM1 (Stott et al. 2006) HadCM3 (Johns et al. 2003) HadGEM2-ES (Jones et al. 2011) Second Generation Canadian Earth System Model (CanESM2; Gillett et al. 2012) Commonwealth Scientific and Industrial Research Organisation Mark, version 3.6.0 (CSIRO Mk3.6.0; Jeffrey et al. 2013)

In this study we investigate how the odds of having a warm year, winter, or summer are changing during years 1–3 with and without human influence in a number of subcontinental “Giorgi” regions (Giorgi and Francisco 2000). We select only regions where our model (HadGEM3-A) has good skill in reproducing warm extremes, as inferred from reliability diagrams (discussed more in section 3). In regions with good reliability we expect certain predictive factors (e.g., SST influence) to largely determine the frequency of extremes. The cases studied in this paper are listed in Table 2. It should be noted that in regions where strong predictive factors are not present, the ACE system can still provide useful attribution assessments as long as the model yields a realistic climatology of extremes and long-term changes in the aspects of climate affecting extremes are well simulated.
In this more general case, an event might have a different climatological frequency in the natural and actual world, which would be independent of the observed SST and not expected to vary considerably from year to year. We define “warm” events as years or seasons when the temperature exceeds its 1961–90 mean value. This relatively low threshold is not an extreme indicator, but our purpose here is to demonstrate the capabilities of ACE and investigate the sensitivity to boundary conditions rather than study specific extreme events. Changes in the chances of events may vary considerably depending on the threshold employed (Christidis et al. 2012). Repeating the analysis with higher thresholds to study specific events would be straightforward and in section 4 we will also consider a more extreme threshold as an example.

3. Human influence on the frequency of warm years and seasons

a. Model evaluation

Reliable attribution assessments require models fit for the purpose, able to represent well the type of event under consideration in the reference region. We evaluate our model using a small ensemble of five long simulations of the actual climate that cover years 1960–2010 (Christidis et al. 2013b). As explained in the previous section, these simulations include both anthropogenic and natural forcings and use HadISST observational data as boundary conditions. Figure 2 shows the regional land temperature time series for the 10 cases investigated in this study constructed with version 4 of the Climatic Research Unit gridded observational temperature dataset (CRUTEM4; Jones et al. 2012). The range from the five model runs is also plotted. Both HadGEM3-A and the observations indicate a warming in all regions and the
FIG. 2. Time series of the annual, summer and winter mean temperature in different Giorgi regions since 1960 calculated with CRUTEM4 (black lines). The green bands illustrate the range from an ensemble of five simulations with HadGEM3-A that cover the period 1960–2010. The model data were masked to include the same spatial coverage as the observations. The 1960–2010 trends corresponding to the observations (black) and the ensemble mean (green) are also noted on each panel. The temperatures are expressed as anomalies relative to the 1961–90 mean.
simulated temperatures are generally consistent with the observations, although in certain cases (e.g., annual MED, and winter ALA) the model tends to produce smaller warming trends.

As in previous work, we use reliability diagrams (Wilks 1995) to examine whether the model-derived probability of a warm event in a region of interest agrees with the observed frequency. In the absence of factors that increase the model predictive skill (e.g., SST), a model would be expected to produce climatological probabilities manifested as a flat line on the reliability diagram. Here, however, we are only interested in cases with good predictive skill and therefore select a cohort of Giorgi regions for which reliability diagrams comprise a line close to the diagonal (Fig. 3). After examining the reliability for annual mean, summer, and winter temperatures in all Giorgi regions, we end up with the 10 cases listed in Table 2, which give the diagrams shown in Fig. 3. The diagrams are constructed as follows. Each of the five 51-yr-long model simulations provides 51 estimates of the regional mean (annual or seasonal) temperature anomaly relative to 1961–90. The forecast probability is computed by checking how many of the five estimates of each hindcast (year or season) are warmer than the upper climatological tercile estimated from the observations. In this way the 51 hindcasts are grouped into five equal-sized probability bins. We finally check the observations to see whether the temperature threshold was indeed exceeded each year, compute the number of observed and nonobserved events for the hindcasts in each probability bin, and use these to estimate the observed frequency as the ratio of the observed and total events. The procedure is discussed in more detail in Christidis et al. (2013b). As shown in Fig. 3, the reliability diagrams are constructed using both CRUTEM4 observations and National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis data (Kalnay et al. 1996). Ideally, we would like to have both longer runs and a larger ensemble to construct better diagrams, but even with the available simulations we have a good indication that in the selected cases the model can skilfully reproduce warm events. It is interesting to note that the ENA region demonstrates good reliability for annual as well as summer and winter temperatures. We find that the SST in the ENA region has a notable influence on the temperature (we estimate a positive correlation of 0.55–065), which appears to play a key role in the frequency of warm events throughout the year.

b. Anthropogenic influence on warm events

We next estimate for each ACE experiment the probability of having a warm event and examine a) how the odds change because of the effect of human influence, b) how they vary from year to year, and c) how sensitive they are to the SSTs prescribed in the natural world simulations. The estimated probabilities are illustrated in Fig. 4. As in previous work, the probabilities are estimated using a statistical extreme distribution if the threshold (i.e., climatological mean temperature) lies at the tails, and a Monte Carlo bootstrap procedure provides an estimate of the sampling uncertainty (Pall et al. 2011; Christidis et al. 2013b). The 5%–95% confidence bands of the probability estimates shown in Fig. 4 are then obtained using a simple order statistics approach.

It is evident that in all cases the overall anthropogenic effect on the climate increases the odds of warm events (red bars in Fig. 4 are always higher than blue bars). In some cases (summer ENA, annual GRN, etc.) the probability of exceeding the warm threshold in the real world is virtually unity in all years, suggesting that the regional anthropogenic warming is so large that every year or season is now expected to be warmer relative to 1961–90. In the colder natural world, however, the odds may decrease significantly as, for example, in the annual GRN case, where the probabilities in year 3 reduce to zero. Cases with high odds of warm events in the natural world are likely to be linked to warm SST in that year, which remains relatively warm even when the anthropogenic effect is removed. For example, in year 1 of the summer GRN case the SST in the region is five standard deviations above the climatological mean. In the following year, the combination of cooler SST (two standard deviations above the mean) and the removal of the anthropogenic SST warming translates to lower probabilities of a warm summer in the natural world.

The example of summer GRN illustrates that as attribution assessments obtained with the ACE methodology are constrained by the prescribed boundary conditions, the resulting probabilities may vary markedly from year to year. In the GRN example warm SST seems to increase the odds of a warm summer, even in the colder natural world. On the contrary, during years or seasons with exceptionally cold SST, the influence of the anthropogenic warming may be relatively smaller and the odds of a warm event may be low as a result, even when human influence is accounted for. This is evident in the winter ALA case, for which the SST is about half a standard deviation colder than climatology in years 2 and 3 and the probabilities of warm events are consequently low (20%–30%) even when the effect of all forcings is included. In the examples shown here the probabilities correspond to a specific temperature threshold, though year-to-year variations could be expected for other thresholds as well. Unlike ACE, previous attribution investigations based on coupled GCMs (e.g., Stott et al.
FIG. 3. Reliability diagrams assessing the model’s skill in representing the frequency of warm events. The panels correspond to the 10 cases investigated in this study and the diagrams are constructed using CRUTEM4 observations (solid lines) and NCEP–NCAR reanalysis data (dashed lines). Warm events are defined as exceedances of the upper climatological (1961–90) tercile estimated from observational or reanalysis data. Proximity to the diagonal (dotted line) indicates good model representation of the event frequency. The inset histograms show the number of hindcasts included in each probability bin. If a bin is underpopulated, there is more uncertainty in the reliability assessment of events within the corresponding probability range.
FIG. 4. Probabilities of warm events in years 1–3 with (red) and without (blue) the effect of human influence on the climate. The panels correspond to the 10 cases investigated in this study and illustrate the best estimate and 5%–95% uncertainty range of the probabilities. There are multiple probability estimates for the natural climate from simulations that employ estimates of the anthropogenic change in the SST from different GCMs (named at the bottom of the figure).
2004) derive probability estimates for the entire range of possible oceanic states, which are not expected to vary considerably from one year to the next. It is therefore essential to clarify what each methodological approach aims to achieve and clearly state the kind of information that the estimated probabilities convey.

The boundary conditions used in simulations of the natural world are found to be a major source of uncertainty in certain regions. In a number of cases the three realizations of the natural world produced with different ΔSST estimates give widely different probabilities. The discrepancy is exacerbated in regions largely influenced by the oceanic state for which the GCM patterns of ΔSST are very different. For example, in year 3 of the summer GRN case the probabilities from the ensemble produced with CanESM2 ΔSST are much lower than the other two ensembles and the relative differences in the probabilities are consistent with the corresponding differences in the ΔSST patterns. The CanESM2 pattern shows more warming in the region, while the CSIRO model includes an area of cooling to the south of Greenland (see annual ΔSST patterns in Fig. 1; the summer patterns show similar features). The probabilities based on the CanESM2 patterns are in fact lower in all cases, as this model indicates larger anthropogenic warming of the oceans around the world. Our findings suggest that the robustness of attribution assessments with the ACE methodology depends crucially upon an accurate representation of the anthropogenic impact on SST. In the next section we attempt to minimize this uncertainty by employing an estimate of ΔSST based on observational data rather than GCMs.

4. Improved representation of the SST in simulations of the natural world

Given the lack of consensus among GCMs on the magnitude and spatial features of the anthropogenic change in SST, we introduce an observational estimate of ΔSST that should provide a better representation. We compute ΔSST on each grid point from the observed trend since the beginning of the HadISST record in 1870 as illustrated in the example of Fig. 5a. We apply a linear fit to the grid point SST time series and approximate ΔSST as the product of the slope and the length of the time series. Provided the time series are long enough (here 142 yr) and assuming there is no significant impact from internal or natural multidecadal variability on the long-term SST trend, the estimated change in the SST is expected to be driven by human influence. A possible disadvantage of this approach could be that the new ΔSST estimate comes from a single observational dataset, whereas the GCM estimates are derived from ensemble means that are less affected by internal variability. However the new representation is based on the trend over 142 years, which is less “noisy” than differences between decadal mean SSTs. The resulting pattern of the anthropogenic change in the annual mean SST is plotted in Fig. 5b. In agreement with the models used in the year 3 experiment (Fig. 1) the observations indicate an overall warming of the oceans, with a global mean magnitude that is similar to the one derived with HadGEM2-ES and CSIRO Mark, version 3.6.0 (CSIRO Mk3.6.0) (~0.5 K), but only half of the CanESM2 estimate. The global change in the annual mean sea ice cover computed with the simple method discussed in section 2 is estimated by the models to be 0.65%–0.95%. The estimate based on observational patterns of the SST change is found to be smaller (0.13%), mainly because of the warming around Antarctica illustrated in Fig. 5b. In certain regions the GCMs may fail to accurately represent the spatial features of the SST change, as in the example of ENA.
illustrated in Fig. 6. The frequency of extremes in this region is found to be sensitive to the oceanic temperature and the probabilities of warm years and seasons derived using different GCM patterns may largely differ (Fig. 4). It is evident in Fig. 6 that none of the modeled estimates of the change in SST due to human influence is in very good agreement with the observed change. We therefore suggest that the most reliable attribution assessment in the region would be the one obtained with the new observational ΔSST estimate.

We carried out a new experiment for year 3 and generated an ensemble of 600 simulations of the natural climate with ΔSST estimated from HadISST data. Figure 7 shows the probability density functions (PDFs) of the regional temperatures in year 3 for the real world and the different versions of the natural world. As in previous work, we correct model biases using the difference between the model and reanalysis climatologies as a correction factor that we apply to all ensembles (Christidis et al. 2013b). More specifically, we estimate the regional mean temperature (annual or seasonal depending on the case) during 1960–2010 from the ensemble mean of the five long model simulations and the reanalysis and use the difference to bias correct the temperatures from simulations of the new experiment. As the same correction is applied to distributions with and without the effect of human influence, their relative separation remains unchanged. Here we assume no model bias in the variance of the distributions, which we find to be a realistic assumption when we evaluate the model against CRUTEM4 observations and reanalysis data. In common with previous studies we bias correct the model using reanalysis data rather than observations because of the poor observational coverage in some regions. Figure 7 shows that while the natural world is always colder, the estimated PDF is highly sensitive to the SST prescribed in the model simulations. The distributions with ΔSST from CanESM2 are in all cases the coldest, as in CanESM2 ensembles the amount of anthropogenic warming extracted from the observations is the largest (Fig. 1). The climatological mean temperature and the temperature in year 3 computed with reanalysis data are also marked on Fig. 7. In certain cases (e.g., annual and seasonal ENA and summer GRN) the expected temperatures in simulations of the actual world are well above the climatological mean. Year 3 is extremely warm in all versions of the natural world (regional temperature is always above the 95% percentile of the natural world distributions) and more consistent with the estimated range of the actual world. There are three cases, however, for which the reanalysis temperature is significantly warmer (above the 95th percentile) than the expected temperature range (annual and summer GRN and summer ALA), even when the effect of anthropogenic forcings is accounted for. We suggest that the interplay among anthropogenic forcings, warm oceans, and internal variability gives rise to the extreme temperatures in these regions.

We finally estimate the fraction of attributable risk (FAR), which measures the fractional change in the probability of a threshold exceedance due to human influence. If $P_1$ and $P_0$ denote the likelihood of exceeding a temperature threshold with and without the anthropogenic effect respectively, then the FAR is simply computed as $1 - (P_0/P_1)$. Values of the FAR closer to unity indicate a greater anthropogenic impact. As mentioned earlier, we estimate $P_0$ and $P_1$ using and an extreme distribution (generalized Pareto), if the threshold lies in the tails. A bootstrap procedure is again employed to generate samples of $P_0$ and $P_1$ from which we compute 10000 estimates of the FAR. The best estimate (50th percentile) and the 5th and 95th percentiles of the FAR are subsequently estimated using order statistics. We investigate the change in the probability of having warm and extreme events in year 3 and define the events using the climatological mean and one standard deviation above the mean as thresholds. Given that the ensemble with the observational estimates of ΔSST provides the most reliable representation of the natural world, we only use this ensemble to derive the probability $P_0$ for the 10 cases examined here. The results are illustrated in Fig. 8. In half of the cases we find that human influence has more than tripled the odds of a warm event and in four cases the increase in the odds is moderate (less than a doubling). However, when a more extreme threshold is considered (Fig. 8b), the effect of anthropogenic forcings greatly increases the likelihood of an event occurrence relative to the natural scenario. A comparison between Figs. 8a and 8b illustrates the increase in the FAR as we move to higher thresholds. We find that extreme events in year 3 are at least 3 times more likely in all cases (best estimate), while in seven cases they are at least 5 times more likely. In conclusion, internal variability appears to play a greater role in the exceedance of moderate thresholds, while more extreme temperatures are most likely to occur in a climate influenced by anthropogenic forcings.

5. Discussion

The new Hadley Centre event attribution system was developed to investigate how certain factors such as human influence on the climate may have affected the frequency of extreme weather and climate events. It is essential that such a system provides high-quality information that can become a good basis for adaptation.
Fig. 6. Anthropogenic change in the annual mean SST in the ENA region estimated (a) with HadISST data and from experiments with the (b) HadGEM2-ES, (c) CanESM2 and (d) CSIRO Mk3.6.0 models.
FIG. 7. Normalized distributions of the regional temperature for year 3 with (red) and without (blue) human influence. The panels correspond to the 10 cases investigated in this study. Different versions of the natural world are illustrated from ensembles that use SST from HadISST data (solid thick blue line), HadGEM2-ES (dashed line), CanESM2 (dash–dotted line), and CSIRO Mk3.6.0 (dotted line). The vertical lines mark the 1961–1990 mean temperature (black) and the year 3 temperature (gray) calculated with NCEP–NCAR reanalysis data.
planning or other decision making. We have so far employed the ACE system for annual assessments of notable events in a year and given the pressing need for timely information (e.g., in the aftermath of a devastating event) we aim to eventually incorporate it into an operational framework to provide assessments in near–real time. Before this happens it is crucial to understand the strengths and limitations of ACE. To that end we have analyzed the three experiments already carried out to demonstrate the system’s capabilities, assess uncertainties associated with the boundary conditions in model simulations, and explore ways of reducing these uncertainties.

A clear indication of an increase in the probability of warm years and seasons as a result of human influence is evident in all the cases considered here. Anthropogenic warming is often so large that exceedance of the climatological mean temperature may now be expected in most years. Nevertheless, precise estimates of the exact change in the odds of warm and extreme events are difficult to obtain in areas heavily influenced by SST because of uncertainties in the model representation of the ocean temperature in simulations of the natural world. In these areas the likelihood of warm and extreme events may vary considerably from year to year depending on the regional SST. We introduce a new way of constructing boundary conditions from observational rather than GCM data that we expect to provide a better representation of ΔSST. Although the GCM estimates employ the mean of ensembles of simulations, which are less influenced by internal variability, they also suffer from model errors and biases. Using a multimodel ensemble mean could reduce the uncertainty from these sources, but might also eliminate features of the SST change that are not correctly represented by all models. The observationally derived ΔSST uses long temperature trends associated with lower uncertainty than the annual or decadal mean temperature. An additional advantage of this new experimental design is that it avoids the need for having several realizations of the natural world generated with ΔSST estimates from several different GCMs, which is computationally expensive. Having an improved representation of the SST change from observations (as in this study) or, possibly, from a multimodel ensemble would be a way forward. Model patterns can also be improved by using observational constraints (Pall et al. 2011). The observational estimate of ΔSST, unlike estimates that rely on GCMs, does not explicitly remove the effect of natural forcings. This, however, is not a major caveat, as the short-lived climatic effect of natural forcings on the 142-yr temperature trend used in the calculation of ΔSST is minimal. Other caveats in our approach are linked to the effects of anthropogenic aerosols and slow-varying modes of variability. Although the linear assumption employed here can adequately describe the temperature response to the dominant anthropogenic forcing from greenhouse gas emissions, it may not be the best way to account for the nonmonotonic effect of anthropogenic aerosol emissions (Booth et al. 2012). Low-frequency modes of internal variability may also have some impact that is not accounted for in estimates obtained either with observations or GCM data.

Finally, it must be stressed that while an accurate SST representation is critical in our methodology, there are several other sources of uncertainty that need to be considered in more detail. One such uncertainty stems from the sea ice representation in simulations of the natural world, which may affect attribution assessments not only in polar regions but also in midlatitudes (Francis and Vavrus 2012). If the simple representation currently adopted is found to be inadequate, improved
methods that give better sea ice estimates need to be developed. Finally, uncertainties in the forcings prescribed in model simulations and other modeling uncertainties may also impact attribution assessments. Attribution systems built on different models developed independently by a number of international research centers will soon enable model intercomparison studies to help assess the impact of modeling uncertainties. While several methodological details remain to be improved and refined, remarkable progress has already been made in this rapidly growing area of research and a number of published studies in recent years have highlighted the potential of event attribution (Peterson et al. 2012, 2013).

Acknowledgments. We are grateful to the reviewers for their constructive comments. This work was supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101).

REFERENCES


