Extreme Value Analysis of Global Temperature Anomalies

Elizabeth Ann Maharaj
Monash University, Melbourne, Australia ann.maharaj@monash.edu

Abstract

Several studies using differing methodologies have revealed that human-made and natural climate forcing variables as well as ocean atmospheric processes contribute to the temperature variation over time, and that the human-made climate forcing is the main contributor to temperature growth. Examination of global land and sea temperatures anomalies for the period 1880 to 2012, from well-known databases, reveal a structural break in the trend from a gentle to a steeper slope from around the mid 1960s. Here an alternative approach that uses generalised extreme value (GEV) analysis is applied to the global temperature anomaly maxima, taking into account the structural break to evaluate the extent of the impact of these climate forcing variables and ocean atmospheric processes on temperature variation. This study reveals that after the structural break, the combined human-made climate forcing variable makes a greater contribution to temperature growth than before it. However, while this contribution is significant, it is smaller than the contribution over the entire record arrived at in other studies. This study also reveals that some of the other climate forcing variables make significant contributions to temperature growth over the entire record. It also exploits a useful feature of the GEV model to produce extreme quantiles of the distribution of annual temperature anomaly maxima, and hence make probability statements concerning long-term temperature changes. This is different from the long-term temperature projections that are made in other studies where fairly complex climate models are used.

Keywords: Temperature anomalies; Climate forcing; Time varying mean; Extreme quantiles.

1. Introduction A large volume of research has established that human-made climate forcing consisting mainly of greenhouse gases and aerosols have contributed to global temperature growth (e.g., Kokic et al., 2013, IPCC2013 and references therein, Estrada et al., 2013, Hansen et al., 2011). The effects of this forcing can either be masked or exaggerated if the naturally occurring climate forcing variables, viz., solar irradiation and stratospheric aerosols (volcanic eruptions) as well as ocean atmospheric processes are not taken into account in the analysis (Estrada et al., 2013). Therefore, taking into account the effects of these factors on temperature variability, will more likely contribute to more realistically assessing the effect of human-made climate forcing variables on temperature growth.

Examination of the global land and sea temperature anomalies show evidence of increasing trends from between the 1960s and 1970s. Estrada et al. (2013). identified structural breaks between the late 1950s and mid-1970s in global and hemispheric temperature anomalies. They also found evidence of a significantly increasing trend in combined greenhouse gases and aerosols climate forcing from around 1960. From their analysis, they implied that by taking into account the non-linear trend in temperature anomalies, all remaining variations in temperatures are stationary and can be explained mostly by non-human factors. Our analysis differs from theirs in that we seek to explain the relationship between human-made factors and temperature changes by simultaneously incorporating the structural break in the global temperature anomalies, the natural climate factors and the atmospheric ocean processes into our model.

Reports from IPCC2013 and Hansen et al. (2011) indicate that the natural climate forcing variables, viz., solar irradiance and stratospheric aerosols (volcanic eruptions which inject aerosols to altitudes of 10-30 km in the stratosphere, where they remain for one to two years reflecting sunlight and cooling the earths surface) do not contribute much to temperature growth. In particular, Hansen et al. (2011) concludes that solar irradiance forcing is small compared to human-made greenhouse gas forcing.

Kokic et al. (2014) reveals from their analysis which assumes a linear trend in the series of temperature anomalies from 1882 to 2010, that human-made climate forcing is the main contributing factor to the increase

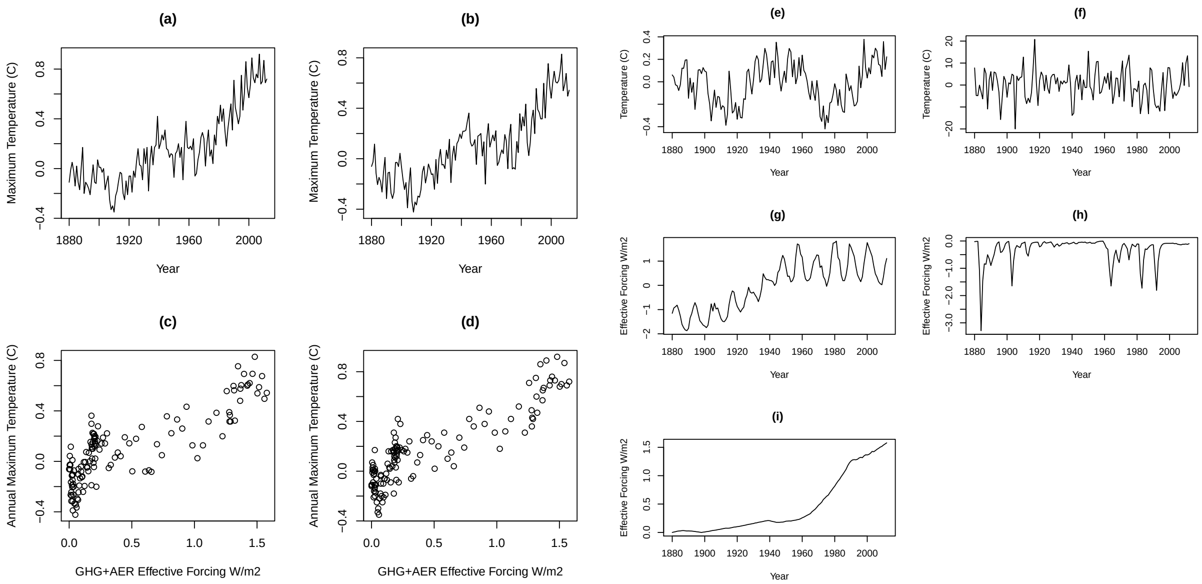
in global temperatures. Studies in econometrics have shown that ignoring the existence of non-linearities in a time series analysis may lead to misleading outcomes (e.g., Peseran and Timmerman, 2004). Our study which takes into account non-linearities in the trend of the temperature anomalies, reveals that while human-made climate forcing contributes positively to global temperature growth before and after identi ed structural breaks, it is not necessarily the main contributing factor over the entire period of the record under consideration. It reveals that solar irradiation and the ocean atmospheric processes, viz., the Atlantic Multidecadal Oscillation Index (AMO) and the Southern Oscillation Index (SOI) contribute at least as much to temperature increases over the entire period of the record as does human-made climate forcing factors.
Using the global land and sea temperature anomalies for 1880 to 2012, from the database of the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS), and from the HADCRUT4 database of the Met O ce Hadley Centre, we t generalised extreme value (GEV) distributions to the series of annual maxima, taking into account, identi ed series-speci c structural breaks. The covariates used in the analysis are, the naturally occurring climate forcing variables of solar irradiance and stratospheric aerosols (volcanic eruptions), human-made climate forcing variables of greenhouse gases and aerosols combined together, and the two ocean atmospheric indices, viz., the AMO and SOI. Sources of these data sets are available from the author on request. The greenhouse gases include well mixed greenhouse gases (carbon dioxide, methane, nitrous oxide, CFCs), ozone and stratospheric water vapour, while the aerosols include re ective tropospheric aerosols, indirect e ect of clouds, snow albedo and black carbon. Aerosol forcing generally has a negative e ect on temperature growth (IPCC2013, Hansen et al. 2011), but when it is combined with greenhouse gas forcing, the overall e ect on temperature growth is positive (Hansen et al. 2011).

2. Examining the Data The NASA series of global land and sea surface temperature anomalies have been developed relative to the reference period 1951-1980 while the reference period for the HadCRUT4 series is 1961-1990. The annual maximum temperature anomalies are shown in Figures 1 (a) and (b) from where it can be observed that in each case, a structural break in the trend is apparent from between the 1960s and 1970s.
The Atlantic Multidecadal Oscillation (AMO) and the Southern Oscillation Index (SOI) are shown in Figures 1 (e) and (f). The AMO is an ocean atmospheric process that has a considerable in uence over the global and northern hemispheric climates (Estrada et al., 2013). Research by several authors including Wu et al. (2011) and Knudsen et al. (2011) have indicated that the AMO tends to exaggerate the warming trend when it is in its positive phase, and tends to mask it in its negative phase. The SOI gives an indication of the development and intensity of $El\ Nino$ or $La\ Nina$ events in the Paci c Ocean. $El\ Nino$ events are associated with the appearance of a warm ocean current o the South American coast and sustained negative Southern Oscillation Index (SOI) values (Australian Bureau of Meteorology website). The Paci c Decadal Oscillation (PDO) the North Atlantic Oscillation (NAO) indices are not included in the analysis because there is a strong relationship between and PDO and the $El\ Nino$ Southern Oscillation (ENSO) (Newman et al., 2003), and because the NAO is strongly associated with the AMO during the winter time (Peings and Magnusdottir, 2014). Correlations between the PDO and SOI and between the AMO and NAO are -0.53 and -0.32, respectively. Figures 1 (g) and (h) shows solar irradiance forcing and stratospheric forcing (volcanic eruptions). Combined greenhouse gases and human-made aerosols forcing which include land use changes are given in Figure 1 (i). It is clear from the graph that there is a slow down in emissions between the early 1930s until about 1960, and then sustained growth until around the mid-1990s. This to some extent coincides with the pattern of growth in the maximum temperature anomalies in Figures 1 (a) and (b). Scatter plots of combined greenhouse gases and aerosols series versus the temperature anomaly maxima in Figures 1 (c) and (d) indicate some degree of non-linearity in this relationship. However, this non-linearity is taken into account in the model tting that follows.

3. Generalised Extreme Value Analysis The GEV distribution has the following form

$$G(x) = \exp\left\{ -\left[1 + \xi\left(\frac{X-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \right\} \tag{1}$$

Figure 1: (a) NASA Maximum Global Temperature Anomalies (b) HadCRUT4 Maximum Global Temperature Anomalies (c)NASA and (d)HadCRUT4: Scatterplot of Temperature Anomaly and Greenhouse Gases + Aersol Forcing (e) Atlantic Multidecadal Oscillation Index (f) Southern Oscillation Index (g) Solar Irradiance (h) Volcanic Eruptions Forcing (i) Greenhouse+Aerosol Gases Forcing



defined on $\{x : 1 + \xi(\frac{x-\mu}{\sigma}) > 0\}$ where $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty$. The three parameters μ, σ and ξ are the location, scale and shape parameters, respectively. When $\xi < 0$, $\xi > 0$ the GEV distribution is the negative Weibull and the Frechet respectively. For $\xi = 0$ it is interpreted as the limit of Equation 1 as $\xi \to 0$. leading to the Gumbel family with distribution:

$$G(x) = \exp\left\{-\exp\left[-\left(\frac{X-\mu}{\sigma}\right)\right]\right\} \tag{2}$$

The implications of a fitted extreme value model are usually made with reference to extreme quantiles. By inversion of the GEV distribution function, the quantile, x_p for a specified exceedance probability *p* is for $\xi \neq 0$

$$x_p = \mu - \frac{\sigma}{\xi}\left[1 - (-\log(1-p)^{-\xi})\right] \tag{3}$$

and for $\xi = 0$

$$x_p = \mu - \sigma \log[-\log(1-p)] \tag{4}$$

x_p is referred to the return level associated with the return period 1/*p*. It is expected to be exceeded by the annual maximum in any particular year with probability *p*. In practice, we often deal with long time series as is the case with the temperature anomolies. Hence, the series of block maxima will no longer be stationary, and adjustments incorporating a time varying mean with *k* covariates, $Y_{1t}, Y_{2t}, ..., Y_{kt}$, can be made to fitting GEV distribution to such series . This is done by replacing the location parameter, μ by $\mu_t = Y_t\beta$, where $Y_t = [1, Y_{1t}, Y_{2t}, ..., Y_{kt}]$ and $\beta = [\beta_1, \beta_2, ...\beta_k]'$ are the regression coefficients. Under the stationary GEV model, the estimators of the parameters are asymptotically normally distributed. For the GEV model with time varying mean, determination of the asymptotic distribution of the estimators of regression coefficients is intractable. Hence, we use bootstrap sampling when the GEV distribution is fitted to the block maxima to obtain percentile intervals which are used to test for significance of the regressions coefficients. Refer to Coles, (2001) for more details about the GEV distribution and Khaliq et al. (2006) for details about the implementation of bootstrap sampling from the GEV distribution.

In order to determine approximate structural break dates in the global series of temperature anomalies, we fitted the GEV distribution to the annual maxima with a time varying mean and two covariates, viz., *time* and $time \times D_t$, where D_t is an index variable set to zero or one.
Hence, if X_t represents the series of temperature anomaly maxima, it is modelled as $GEV(\hat{\mu}_t, \hat{\sigma}, \hat{\xi})$, where

$$\hat{\mu}_t = b_0 + b_1 time + b_2(time \times D_t) \tag{5}$$

is the time varying mean or location estimator, and $\hat{\sigma}$ and $\hat{\xi}$ are the scale and shape estimators, respectively. The variable *time* is coded from 1 to 133 to cover the years 1880 to 2012, b_0 is the intercept while b_1 and b_2 are the slopes associated with the *time* and $time \times D_t$, respectively. D_t, the index variable is set to zero or one. In order to determine if the trend significantly increases after a particular date (year), bootstrap tests are conducted with D_t being set in turn to 0 for 1880, 1 for 1881 to 2012, 0 for 1880-1881, 1 for 1882-2012, ... , 0 for 1880-2011, 1 for 2012. For the NASA series, it was determined that there is a significant increase in the trend in the series of temperature anomaly maxima at the 1% level from 1962 to 1963, while for the HadCRUT4 series, it was determined to be from 1970 to 1971.
Using the 1963 and 1971 break dates, we fitted the GEV distribution to NASA and HadCRUT4 temperature anomaly maxima, respectively, by including the interaction of the human-made forcing variables with an index variable, D_t. This was done by setting D_t to zero for 1880 to 1962 and to one 1963 to 2012 to coincide with the 1963 break of the NASA series, and by setting D_t to zero for 1880 to 1970, and to one for 1971 to 2012 to coincide with the 1971 break for the HadCRUT4 series. Using the interaction variable to take into account the structural break in the temperature anomalies addresses the non-linearity in the relationship between temperature anomaly maxima and the combined GHG+ARE forcing as observed in Figures 1 (c) and (d). If X_t represents the series of temperature anomaly maxima, then it is modelled as $GEV(\hat{\mu}_{t1}, \hat{\sigma}_1, \hat{\xi}_1)$, where

$$\hat{\mu}_{t1} = b_0 + b_1(GHG + AER) + b_2((GHG + AER) \times D_t) + b_3 SOL + b_4 AMO + b_5 SOI + b_6 VOL \tag{6}$$

is the time varying mean estimator, and $\hat{\sigma}_1$ and $\hat{\xi}_1$ are the scale and shape estimators, respectively. b_0 is the intercept and b_1 to b_6 are the slopes associated with the covariates. The estimated coefficients of the interaction variable determines the contribution of the human-made climate forcing variable, GHG+AER, to temperature growth before and after the break date. The maximum likelihood estimates (MLE) and standard errors (SE) of the GEV models fitted to the temperature anomaly maxima of the NASA and HadCRUT4 series for the 1963 and 1971 break dates, respectively, are given in Table 1. The adequacy of the fitted models

Table 1: Maximum liklihood estimates and standard errors of fitted GEV models

	NASA			HadCRUT4		
	MLE	**SE**		**MLE**	**SE**	
Intercept	0.005	0.033		-0.065	0.035	
GHG+AER	0.093	0.055	*	0.049	0.060	
(GHG+AER)*Dt	0.156	0.068	*	0.161	0.075	**
SOL	0.072	0.015	***	0.089	0.016	***
AMO	0.157	0.018	***	0.173	0.019	***
SOI	-0.040	0.020	**	-0.059	0.019	***
VOL	-0.005	0.009		0.003	0.010	
Scale	0.097	0.007	***	0.100	0.007	***
Shape	-0.175	0.074		-0.197	0.059	*

can be assessed by examining diagnostic plots, viz., probability and quantile plots of the standardised version of the fitted data (Coles, 2001). Diagnostic plots of the fitted models to the NASA and HadCRUT4 series (not shown in the paper but can be obtained from the author on request) indicate that GEV model appears to fit both series reasonably well.

In order to assess the significance or otherwise of the coefficients of the climate forcing variables, we obtain bootstrap percentile intervals. In Table 1, *, **, ***, imply significance at the 10%, 5% and 1% levels, respectively, i.e., tests of significance are based on 90%, 95% and 99% bootstrap percentile intervals. From Table 1, we make the following observations about the covariates in the estimated time varying mean parameter of the GEV model: The AMO makes a significant positive contribution to temperature growth at the 1% level during its positive phases compared to its negative phases for both series. The SOI makes a negative contribution to the temperature growth during its positive phases than during its negative phases and this contribution is significant at the 5% level for the NASA series and 1% level for the HadCRUT4 series. Solar irradiation contributes significantly to temperature growth at the 1% level for both series. The effect of stratospheric aerosols (VOL) on the temperature trend is negative for the NASA series and positive for the HadCRUT4 series; in both cases it is not significantly different from zero and hence it does not make a significant contribution to temperature growth. For each series, the combined greenhouse gases and aerosols forcing contribute positively to temperature growth before and after the identified structural break, with this contribution being greater after the break than before it. For the NASA series, this contribution is significant before and after the break at the 10% level, whereas for the HadCRUT4 series, it is significant at the 5% level, only after the break. For both series, the scale parameter is significantly different from zero at the 1% level, while the shape parameter is significantly different from zero for the HadCRIT4 series at the 1% level. We assessed the robustness of the model by shifting break dates to years before and after the identified breaks dates of 1963 and 1971 for the NASA and HadCRUT4 series, respectively. The results of these robustness checks including the goodness of fit graphs can be obtained for the author on request. For the NASA series, we found that the model was robust in terms of the significance of the coefficient of the interaction variable, $(GHG + AER) * D_t$, a year before and a year after the initially identified break date, whereas for the HadCRUT4 series, it was a robust for two years before the identified break date and four years after it. Hence, it appears that the significant contribution of the combined greenhouse gases and aerosols forcing to temperature growth co-insides with break dates between 1962 and 1964 for the NASA series, and between 1969 to 1975 for the HadCRUT4 series, and that this contribution is greater after each of these these break dates than before them. However, the highest contribution occurred after the initially identified break date for each series, i.e., 1963 for the NASA series, and 1971 for the HadCRUT4 series.

4. Extreme Quantiles A useful feature of GEV analysis is being able to obtain extreme quantiles, also referred to as returns. Extreme quantiles are obtained by substituting the estimated mean, for a particular year of the given record, together with the estimated scale and shape into Equation 3. Based on the NASA data, temperature anomaly maxima of 1.03^oC, 1.06^oC and 1.09^oC are expected to be exceeded on average once every 25, 50 and 100 years, respectively. Based on the HadCRUT4 data, temperature anomaly maxima of 0.90^oC, 0.94^oC and 0.97^oC are expected to be exceeded on average once every 25, 50 and 100 years, respectively. Given that the largest anomaly maxima for the years under consideration, i.e., 1880 to 2012, is 0.92^oC for the NASA series and for 0.83^oC, for the HadCRUT4 series, these returns appear to be realistic.

5. Conclusions The main finding of this study is that while human-made climate forcing, most of which is comprised of greenhouse gas forcing, makes a positive contribution to temperature growth, it contributes significantly to temperature growth from around the mid-1960s, at a greater rate than it did before that, even when the contribution of the natural climate forcing elements and oceanic processes are taken into account. The positive contribution of human-made climate forcing to temperature growth is in keeping with the finding of other studies, where different methodologies are used (e.g., Kokic et al. 2014, IPCC2013, Estrada et al. (2013), Hansen et al. (2011). So in essence, we have demonstrated this consistency of the finding in climate change research, using yet another methodology, viz., GEV analysis. However, where our analysis differs from that of these other studies is that taking into account the non-linearity in the relationship between temperature anomalies and the human-made climate forcing variable, i.e., the structural break in the temperature anomaly maxima, leads to a smaller yet still significant contribution of the human-made climate forcing variable to temperature growth. It reveals that for every unit increase in effective human-made climate forcing, while keeping the other factors constant, the annual global temperature anomaly maximum increases by approximately 0.25^oC for the NASA series and by 0.21^oC for the HadCRUT4, respectively, after the structural break, whereas before the structural break this increase is approximately 0.09^oC and 0.05^oC, respectively.

The analysis by Kokic et al., (2014) which includes similar covariates to the ones we used except for the AMO, reveals that for every unit increase in e ective human made climate forcing, while keeping the other factors constant, the average annual global temperature anomaly increases by approximately 0.40^{o}C between 1882 and 2010 using two other data sets, similar to the ones used in our analysis. However, they did not take into account the non-linearity in the relationship between temperature anomalies and the human-made climate forcing variable. Our analysis also reveals the signi cance of the contribution of solar irradiance, and the atmospheric ocean processes, viz., the AMO and SOI in their positive and negative phases, respectively, to global temperature growth. However, studies from the above-mentioned authors including IPC2013 indicate that solar irradiance does not make a signi cant contribution to global temperature growth. Hence, this is another point of di erence with the outcome of our study and that of these other studies. Finally, a useful feature coming out of our statistical study based on GEV analysis, which is a further point of di erence from other statistical studies, is the the ability GEV analysis to make probability statements concerning long-term temperature changes. All other studies where long-term climate predictions that are made, are based on complex climate models.

References

[1] Australian Bureau of Meteorology, http://www.bom.gov.au/climate/glossary/soi.shtml

[2] Coles, S. (2001). An introduction to statistical modeling of extreme values Springer-Verlag: London.

[3] Estrada, F., Perron, P., Mart$i$nez-L$o$pez, B. (2013). Statistically derived contributions of diverse human in uences to twentieth-century temperature changes, Nature Geoscience, 6, 1050-1055.

[4] Hansen, J., Sato, M., Kharecha, P., von Schuckmann, K. (2011) Earth's energy imbalance and implications, Atmos.Chem.Phys., 11, 14321-13449.

[5] IPCC2013 (2013). Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[6] Khaliq, M.N., Ouarda, T.B.M.J., Ondo, J.C., Gachon, P., Bob$e$e, B. (2006) Frequency analysis of sequence of dependent and/or non-stationary hydro-meteorological observations: A review, Journal of Hydrology, 329, 534-552.

[7] Kaufmann R. K., Stern, D.I. (1997). Evidence of human in uences on climate from hemispheric temperature relations, Nature, 388, 38-44.

[8] Kelsey, D., Kozhan., R., Pang, W. (2010). Spurious Regressions of Stable AR(p) Processes with Structural Breaks, Working Papers wp09-04, Finance Group, Warwick Business School. University of Warwick).

[9] Knudsen, M. F., Seidenkrantz, M.S.,Jacobsen, B.H., Kuijpers, A. (2011). Tracking the Atlantic Multidecadal Oscillation though the last 8000 years, Nature Comm. 2, 178.

[10] Kokic, P., Crimp, S., Howden, M. (2014). A probabilistic analysis of human in uence on recent record global mean temperature changes, Climate Risk Management, 3, 1-12.

[11] Newman, M., Compo, G. P., Alexander, M.A. (2003). ENSO-Forced Variability of the Paci c Decadal Oscillation, Journal of Climate Letters, 16, 23.

[12] Peings J., Magnusdottir,G., Response of the Wintertime Northern Hemisphere Atmospheric Circulation to Current and Projected Arctic Sea Ice Decline: A Numerical Study with CAM5, Journal of Climate, 27, 3853-3857 (2014).

[13] Peseran, M. H., Timmerman, A. (2004). How costly is it to ignore breaks when forecasting the direction of a time series? International Journal of Forecasting 20, 411425.

[14] Wu, Z., Huang, N.E.,Wallace, J.M., Smoloak, B.V. Chen. X. (2011). On the time-varying trend in global-mean surface temperature, Clim Dyn. 37, 759-773.