**Using IMAGE, a multi-site, multivariate stochastic weather generator, to model European extreme weather under climate change**

Robbie M Parks1, Steven Hardwick1, Nathan Sparks1, Ralf Toumi1\*

1Space and Atmospheric Physics, Imperial College London, London, United Kingdom

\*corresponding author:

Ralf Toumi

Imperial College London

Huxley Building

South Kensington Campus

London

SW7 2AZ

E-mail: [r.toumi@imperial.ac.uk](mailto:r.toumi@imperial.ac.uk)

Tel: +44 (0)20 7594 7668

**Abstract**

Heat waves negatively affect society through increased mortality and economic losses. As a result of rising global temperatures, the number of extreme heat events is set to increase. Governments and businesses need to prepare for this increased risk of heat waves, however, modelling the frequency and intensity of rare events can be difficult using dynamical climate models as they are too computationally expensive to be run for the number of simulation years required. Here, we present an approach to this problem using the Imperial College Weather Generator (IMAGE), a massively multi-site, multivariate stochastic weather generator. IMAGE was provided thirty year time slices from regional climate model output produced as part of the EURO-CORDEX experiment and was used to generate 10,000 years of synthetic data for two different future time periods (2021-2050 and 2071-2100) under two different climate scenarios (RCPs 4.5 and 8.5), as well as for a historical period (1971-2000). IMAGE successfully reproduced the mean temperatures and spatial correlation of the input data. IMAGE reproduced return values in the input dataset for heat waves as measured by both intensity and duration with reasonable accuracy. The 10,000 year IMAGE simulations predict increases in the intensity and duration of rare heat waves under climate change, with mean daily maximum temperature being 3.92 °C warmer for Europe-wide heat waves with a return period of 100 years under the RCP 8.5 scenario.

**Introduction**

Extreme heat events have negative impacts on society through increased mortality (D’Ippoliti et al, 2010; Gasparrini and Armstrong, 2011), agricultural losses (Ciais et al, 2005) and damage to property and infrastructure (Kovats et al, 2014). The major European heat wave of 2003 was estimated to have caused more than 14,000 additional deaths in France and losses of over $10 billion across Europe, due to health impacts, forest fires and damage to property, livestock and crops (García-Herrera et al, 2010). Global mean surface temperature is increasing as a result of anthropogenic emissions of greenhouse gases. Alongside global mean warming, the magnitudes of extreme heat events at individual locations around the world have increased since 1900 (Donat et al, 2013). Anthropogenic climate change was found to have increased the risk of mortality in Paris and London by 70% and 20%, respectively, during the heat wave of 2003 (Mitchell et al, 2016). Therefore, predicting future risk due to heat waves is of vital importance to society.

Stochastic weather generators produce long, realistic time series of weather variables that can be used for risk planning, by both commercial organisations, such as insurance companies, and public bodies. Stochastic weather models require substantially less processing power than dynamical climate models and so can be used to generate many thousands of years of synthetic weather data to allow the magnitude and frequency of extreme events to be estimated. The most basic stochastic weather generators focus only on producing time series at a single site, however, some applications require spatially correlated weather time series at multiple locations. [[[Comparison of other weather generators here]]]

Here, we use the Imperial College Weather Generator (IMAGE) to assess future risk of heat waves in Europe under climate change. IMAGE is a multisite, multivariate stochastic weather generator that can quickly simulate thousands of years of synthetic weather data. IMAGE has previously been shown to perform well in comparison to other multivariate, multisite weather generators and in particular has demonstrated its ability to realistically simulate climate extremes (Sparks et al, 2017).

In this study, we made one key alteration to the methodology of IMAGE that improves its ability to preserve the spatial correlation of the input data. Accurate simulation of spatial correlation is very important, because as well as predicting changes in the frequency and magnitude of heat waves, the output of IMAGE can be used to drive, for example, mortality models that have been constructed using district level data (e.g. Bennett et al, 2014). This means governments can plan for realistic heat wave scenarios where some regions are more affected than others, rather than assuming a uniform impact across a country. This spatial information is also useful for insurers, who may have risk spread across different countries. A heat wave across a large part of Europe will have much a larger impact on their risk portfolio than a heat wave in only one country. By using regional climate projections produced by the EURO-CORDEX project (Jacob et al, 2014) as input to IMAGE, we demonstrate its potential use in estimating the future risk of extreme events under climate change.

**Model Description**

IMAGE is described fully in Sparks et al (2017), and only a brief description is included here, with detailed explanation only included for parts of the model that have been changed significantly for this study.

All variables in IMAGE are modelled as latent Gaussian variables. At the start of simulation each variable is transformed using a normal quantile transformation such that it has a normal distribution. In this study, these transforms are performed separately for each calendar month to allow for changes in the distribution of temperature from month to month. Once transformed, an autoregressive lag 1 model of the form

(1)

is fitted separately to each month of input data for each variable at each site, where *cs* is a constant, *αs* is referred to as the memory parameter and *ϵs* is a noise term. These three parameters are each, in turn, modelled as latent Gaussian variables and are transformed such that each parameter has a normal distribution for each variable at each site for each calendar month.

Synthetic time series are simulated for each variable at each site by first generating correlated values of *cs* and*αs* for each month by sampling from a multivariate normal distribution. This process requires the decomposition of the covariance matrix of the autoregressive parameters, *Σ*, to a matrix *C* such that . In Sparks et al (2017) this is achieved using empirical orthogonal function decomposition, however, in this study we instead use Cholesky decomposition, which produces the same results but is computationally quicker. In general, *Σ* may not be semi-positive definite, which is required when sampling from the multivariate normal distribution, therefore the nearest semi-positive definite matrix to *Σ* is computed using the method of Higham (1988). Parameters are generated simultaneously for all twelve months in one simulated year, such that correlations between months in the same year are accurately simulated, as well as spatial correlation between sites. As well as simulating monthly parameters, the noise terms *ϵs* are simulated daily for each variable at each site, once again by sampling from a multivariate normal distribution. Daily values for each variable at each site can then be simulated using Eq. (1).

After simulation, variables are transformed back to their original distributions using an inverse normal quantile transformation. The pairwise Pearson’s correlation coefficient of time series of variables at different sites are calculated for the simulated data and compared to the correlation coefficients of the input data. As described in Sparks et al (2017), the original version of IMAGE tended to systematically under-simulate the observed spatial correlations. Therefore, in this study, once one simulation run is complete, the covariance matrix that is used to generate the daily noise terms, *ϵs*, is adjusted by applying a correction term equal to the difference between the observed correlation and the simulated correlation for each pair of sites. The simulation of *ϵs* is then re-run and this cycle is iterated until a satisfactorily small error in the simulated pairwise correlations is achieved. After testing, it was found that ten iterations were sufficient to reduce errors in correlations to an acceptable level.

The modelling process is summarised in Figure 1.

**Method**

Regional climate projections generated for Europe as part of the EURO-CORDEX project (Jacob et al, 2014) were used as input data for IMAGE in this study. EURO-CORDEX projections were produced by downscaling various different CMIP5 general circulation model projections using a selection of regional climate models. As the purpose of this study was to demonstrate the potential for using IMAGE to generate long time series of weather under future climate scenarios, only one member of the EURO-CORDEX ensemble was selected to provide input data for IMAGE. So as to produce results that are broadly consistent with the ensemble mean, the ensemble member that most accurately reproduced the ensemble mean warming trend across Europe over the period 2006-2100 under the RCP 4.5 scenario was selected. This was found to be a simulation using REMO regional climate model (Jacob et al, 2012) to downscale the Max Planck Institute Earth System Model (MPI-ESM) from CMIP5 ensemble r2i1p1.

Thirty year time slices of dynamical model data were used as input to IMAGE, and it was assumed that these data represented a stationary climate. From the historical run, the period 1971-2000 inclusive was selected. For both RCP 4.5 and RCP 8.5 future scenarios, the time periods 2021-2050 and 2071-2100 inclusive were selected. Two variables were simulated, the daily maximum air temperature, Tmax, and the daily apparent temperature, Tapp, which has been found to be a useful predictor of heat wave mortality in some circumstances (Barnett et al, 2010; O’Neill et al, 2003) and which is defined as

(2)

where Tmean is the daily mean air temperature and Td is the dew point temperature. While the methodology of IMAGE places no restrictions on the number of sites that can be simulated at once, computational limitations meant that the REMO model output data was re-gridded to 0.88° resolution. In total, 10,000 years’ worth of daily sequences of Tmax and Tapp were generated for 805 grid cells for each thirty year time slice of input data.

For the purposes of examining the magnitude of severe heat waves under climate change, two metrics were chosen. The first metric was the annual peak mean daily maximum air temperature over three consecutive days between May and September and was chosen to represent the intensity of heat waves. The second was the number of consecutive days with daily maximum air temperature exceeding the 99th percentile of daily maximum air temperature during May to September of the control period, 1971-2000, and was chosen to represent the duration of heat waves. Heat wave intensity and duration were examined across the whole domain and also in selected countries (Spain, France, Germany, Italy, United Kingdom). In each case the mean daily maximum air temperature across all grid cells within the selected region was used in the two heat wave metrics described above.

**Results**

Model validation was performed by comparing the IMAGE simulation for the period 1971-2000 with the CORDEX input data used for that period. IMAGE accurately reproduced daily means of both Tmax and Tapp in each month (Fig. 2) (mean bias, Tmax and Tapp: < 0.001 °C; RMSE, Tmax = 0.02 °C, Tapp: 0.02 °C). IMAGE also accurately simulated the daily standard deviations of both Tmax and Tapp (Fig. 3) (mean bias, Tmax: -0.05 °C, Tapp: -0.05; RMSE, Tmax: 0.05 °C, Tapp: 0.05 °C). However, IMAGE consistently overestimated the monthly standard deviations of both Tmax and Tapp (Fig. 4) (mean bias, Tmax: 0.67 °C, Tapp: 0.70 °C; RMSE, Tmax: 0.67 °C, Tapp: 0.70 °C). This means that year-to-year variability in the temperatures produced by IMAGE is greater than that in the input data, however the day-to-day variability within a given year is accurately simulated. IMAGE also accurately reproduced the spatial correlation of both Tmax and Tapp as well as the spatial cross-correlation between Tmax and Tapp (Fig. 5), with the mean bias in the simulated Pearson correlation coefficient being negative with magnitude less than 0.003 in all cases and the RMSE in all cases being less than 0.01. Autocorrelation was systematically over-simulated by IMAGE for both Tmax and Tapp for lags of one, three and five days (Fig. 6) (mean biases: Tmax, lag 1 day: 0.06; 3 days: 0.14; 5 days: 0.15; Tapp, lag 1 day: 0.18; 3 days: 0.18; 5 days: 0.19).

IMAGE tended to under-simulate the intensity of heat waves when the mean daily maximum air temperature across the whole domain was considered (Fig. 7). Events with return periods of 2, 5, 10 and 30 years were cooler in the IMAGE simulation than in the input CORDEX data (biases: 2 year return period: -0.61 °C; 5 year return period: -0.81 °C; 10 year return period: -0.61 °C; 30 year return period: -0.48 °C). IMAGE performed similarly when considering individual countries within the domain. In all five countries considered the intensity of events with a return period of 2 years were cooler in IMAGE than in the input CORDEX data, with the mean bias across all countries being -0.80 °C. Rarer events with longer return periods had smaller mean biases and lower RMSEs, with a general tendency for IMAGE to produce cooler heat waves than in the input CORDEX data (5 year return period: mean bias: -0.42 °C, RMSE: 0.55 °C; 10 year return period: mean bias: -0.32 °C, RMSE: 0.53 °C; 30 year return period: mean bias: -0.52 °C, RMSE: 0.52 °C). In all cases, the 200 year return period event simulated by IMAGE was warmer than any event in the input CORDEX data.

Similarly to the results for heat wave intensity, IMAGE under-simulated the duration of heat waves occurring across the whole domain (Fig. 8) (biases: 2 year return period: -1 day; 5 year return period: -2 days; 10 year return period: -3 days; 30 year return period: -4 days). However, for individual countries IMAGE generally performed better. There was a slight tendency for IMAGE to under-simulate the duration of heat waves in the five countries considered, however, the mean biases were very small (mean biases: 2 year return period: -0.6 days; 5 year return period: -0.4 days; 10 year return period: 0 days; 30 year return period: 1.6 days). As with heat wave intensity, in all cases the 200 year return period event simulated by IMAGE had longer duration than any event in the input CORDEX data.

IMAGE projected significant increases in the intensity and duration of rare heat waves under climate change (Figs. 9, 10, 11, 12). For example, heat waves with a 100 year return period as measured by intensity across the whole domain were warmer in 2071-2100 by 1.52 °C compared to 1971-2000 under the RCP 4.5 scenario and warmer by 3.92 °C under the RCP 8.5 scenario. Similar trends were seen for individual countries, with Spain being the worst hit, with the projected 100 year return period heat wave as measured by intensity being 5.88 °C warmer in 2071-2100 compared to 1971-2000 under the RCP 8.5 scenario. Heat waves with a 100 year return period as measured by duration across the whole domain lasted 7 days in the period 1971-2000, while in 2071-2100 they lasted 27 days under the RCP 4.5 scenario and 62 days under the RCP 8.5 scenario. The same was true for each of the five individual countries considered, with Spain once again being the worst hit, with the 100 year return period heat wave as measured by duration lasting for 9 days in 1971-2000 compared to 21 days in 2071-2100 under the RCP 4.5 scenario and 60 days under the RCP 8.5 scenario.

Maps of temperature across Europe as simulated by IMAGE during nine heat waves in France with return periods close to 100 years, as measured by intensity (Fig. 13), show that IMAGE can generate very different spatial patterns of temperature for events of the same magnitude. In each case the mean daily maximum temperature for three consecutive days across France is 32.4 °C, but across the United Kingdom, a neighbouring country, during the same nine events, the mean daily maximum temperature varies between 17.8 and 25.3 °C. In Greece, on the other side of Europe, during the same nine events the mean daily maximum temperature varies between 24.6 and 31.4 °C.

Jacob et al (2014) mapped the increase in the number of heat waves projected to occur under climate change in the CORDEX simulations. Reproduced here (Fig. 14) is their map of the ensemble mean number of additional heat waves in 2071-2100 compared to 1971-2000 under the RCP 8.5 scenario, where a heat wave is defined as more than three consecutive days with the daily maximum temperature exceeding the 99th percentile of the daily maximum temperature in the months from May to October during the period 1971-2000. For comparison, the same map was produced for the one ensemble member being used in this study (Fig. 15a), which shows that the trend projected is broadly similar to the ensemble mean. The overall distribution produced by IMAGE is similar (Fig. 15 b), with more heat waves expected to occur in Southern Europe than in Northern Europe, however there is a noticeable bias between IMAGE’s simulations and those of the CORDEX input data (Fig. 15 c). IMAGE simulates more heat waves in Germany and Eastern Europe and less heat waves in Southern Europe, the UK and Scandinavia. Across the whole domain, IMAGE simulated a mean of 1.6 heat waves less than in the CORDEX input data over 30 years, while the RMSE was 5.5 heat waves.

**Discussion**

IMAGE successfully reproduced the daily means and daily standard deviation of the input data. The simulated standard deviation of the monthly means had a significant positive bias during all months. As discussed in Sparks et al (2017) this is, in general, the opposite problem to that of other stochastic weather generators, which tend to have too little variability from year to year. This issue could be resolved by relaxing the simulated autoregressive parameters towards the observed climatological mean each month, but doing this results in less accurate simulation of extreme events, which are the main focus of this study, therefore the biased monthly standard deviations were considered an acceptable trade-off.

One key difference between the statistics presented in this study and those in Sparks et al (2017) was the accurate representation of the spatial correlation of all variables. This significant improvement was a result of the novel iterative approach used when generating the daily residuals during the simulation. This is very important for future planning. While a single-site stochastic weather generator could be used to estimate return values on a Europe-wide basis, this would provide no spatial information on which areas are more or less severely impacted. By generating daily spatial temperature fields, IMAGE allows the correlated risk within the domain to be calculated. This is illustrated by the maps in Figure 13, which show that for events with very similar mean daily maximum temperatures in France, there can be a large difference in temperatures in other parts of Europe. An EU-wide response to a major heat wave could involve sharing of resources across national borders. Therefore, preparing such a plan would require the sort of information provided by IMAGE on the risk of simultaneous heat waves in different countries. Similarly, insurers would be interested in the risk of simultaneous extreme events across Europe when calculating their full portfolio risk. Additionally, the same mean daily maximum temperature throughout the domain can correspond to very different spatial distributions of temperature which can have different impacts on, for example, human mortality. As IMAGE generates these different possible spatial distributions they can be used as input to mortality models and allow these differences to be investigated. While the error in spatial correlation was very small, there was a systematic bias in the simulated autocorrelation. The root cause of this bias is not clear and will be a subject of future investigation.

IMAGE generally performed well when simulating the intensity and duration of heat waves for individual countries, but tended to under-simulate the severity of heat waves when compared to the input data. This is slightly surprising given that IMAGE tends to over-simulate autocorrelation, which one might assume would lead to longer and more intense heat waves than in the input data set. IMAGE’s simulations are based purely on the mean correlations over the entire input data set so these results suggest that spatial correlation and/or autocorrelation behave differently during extreme events than their average behaviour suggests. If, for example, autocorrelation tends to be much higher during heat waves than during periods of normal temperatures, then it may be that IMAGE is under-simulating the autocorrelation and thereby simulating heat waves that are too short and less intense. Similarly, if extreme heat events have different spatial footprints to days with normal temperatures, then again IMAGE would not be able to reproduce these spatial patterns correctly and could end up underestimating the frequency of severe domain-wide or countrywide heat waves. An alternative approach that could be worth investigating would be to model only extreme events rather than modelling 10,000 continuous years of data. A similar approach was used with some success by Youngman and Stephenson (2016) when modelling extreme wind events in Europe.

The 10,000 year simulations produced by IMAGE project that rare heat waves will become more extreme under climate change. The absolute temperature values simulated in this study are possibly not accurate, as the input CORDEX data used had not been bias-corrected, however the size of the projected changes can be analysed. When considering the whole domain, the intensity of the 100 year return period heat wave increases by 1.52 and 3.92 °C by 2071-2100 compared to 1971-2000 under the RCP 4.5 and 8.5 scenarios, respectively. Such increases will clearly have consequences for human health and mortality as well as the economic impact of agricultural losses and damage to infrastructure and property. When considering individual countries, the changes are largest in Southern Europe, which is in agreement with the findings of Jacob et al (2014).

While the results described here are only for one ensemble member of the EURO-CORDEX experiment, the particular ensemble member was selected as it most closely matched the ensemble mean warming trend under the RCP 4.5 scenario. It also simulated a very similar number of additional heat waves in 2071-2100 compared to 1971-2000 under the RCP 8.5 scenario to the ensemble mean (Figs. 14 and 15). Therefore, we can have some confidence that the results presented in this study for the changes in magnitude of heat waves under climate change could be quite close to the ensemble mean projections.

**Conclusion**

We have demonstrated the use of IMAGE for assessing the risk of extreme events under climate change. Short time slices of dynamical climate model output can be extended to very long time series so that the return values of rare events can be estimated. IMAGE’s accurate reproduction of spatial correlation means that the impact of heat waves on processes that depend on the spatial distribution of temperature, such as human mortality or agricultural loss, can be investigated using data simulated by IMAGE. Rare heat waves across Europe will become more severe and have longer durations under climate change, with Southern Europe projected to experience larger changes than Northern Europe.