

INCORPORATING CLIMATE CHANGE PREDICTIONS IN THE ANALYSIS OF WEATHER-BASED UNCERTAINTY

Parag Rastogi and Marilyne Andersen

Interdisciplinary Laboratory of Performance-Integrated Design (LIPID)

École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne, Switzerland

parag.rastogi@epfl.ch

ABSTRACT

This paper proposes randomly-generated synthetic time series incorporating climate change forecasts to quantify the variation in energy simulation due to weather inputs, i.e., a Monte Carlo analysis for uncertainty and sensitivity quantification. The method is based on the use of a small sample (e.g., a typical year) and can generate any numbers of years rapidly. Our work builds on previous work that has raised the need for viable complements to the currently-standard typical or reference years for simulation, and which identified the chief components of weather time series. While we make no special efforts to reproduce either extreme or average temperature, the sheer number of draws ensures both are seen with either the same or higher probability as recent recorded data.

INTRODUCTION

The analysis of uncertainty in building simulation is related to the need for risk-conscious design. Existing studies have largely focused on analysing the effect of uncertainty in material inputs (a kind of epistemic uncertainty), and variations due to occupant behaviour (a type of aleatory uncertainty). A few have focused on examining the impacts of climate change and the effect of the uncertainties inherent in simulation based on ‘future weather’ (e.g., Belcher et al. 2005; Chinazzo et al. 2015a; Crawley 2008; Jentsch et al. 2008; Kershaw et al. 2011; Wilde et al. 2008). These uncertainties in the weather input arise due to modelling assumptions (simplifications of physical phenomena, skipping phenomena that are not well understood), incomplete records (to calibrate climate models), ‘downscaling’ (where global circulation models have to be ‘scaled’ down to a region of interest), among other sources. Kershaw et al. (2011) argue that using a single typical or reference file is conceptually and computationally far simpler than working with several files, each of which have some probability of occurring. They point out that while the original advantages of reducing simulation time by incorporating smaller weather files should now be irrelevant, the increasing complexity of building simulation codes has negated much of the gain in computational speed. In any case, typical files, of any sort,

cannot be used to assess risk.

To simulate the future performance of a building, i.e., an explicit estimate of some performance parameter conditional on physically viable future projections, a ‘future weather file’ is needed. Belcher et al. (2005) proposed ‘morphing’, a simple solution that can be easily implemented in the context of building simulation, since it only requires one of three operations: addition (shifting), multiplication (linear stretching), and a combination of the two (shift and stretch). Shifting is applied to those variables for which an absolute change of mean is available in climate change forecasts. Stretching works when the change to mean or variance is given as a fractional change. Finally, the combination is used when both the mean and variance of a variable need to be changed. For example, if the forecast includes a change of minimum and maximum temperatures in addition to a change of mean temperatures. Belcher et al. (ibid.) demonstrated their method for three cities in the United Kingdom. They demonstrated the agreement of future heating degree day values calculated using their ‘morphed’ Test Reference Year (TRY) and Design Summer Year (DSY) files, with those calculated directly from the UKCIP02 report itself (the forecasts on which the morphed files were based). Kershaw et al. (2011) used the latest future weather generator from the UK, the UKCP09, which is based on a future rainfall generator model. The baseline climate, like most generators and projections, is 1961-1990. Upon calibration, “change factors [were] applied to [recorded data to] generate the future precipitation”. All “... other variable [were] created using mathematical and statistical relationships with daily precipitation and the previous day’s weather” (ibid.). The UKCP09 generator outputs 100 runs of 30 years each, from which the authors constructed 100 reference years. Eames et al. (2011) used the percentiles of monthly mean Dry Bulb Temperature (TDB) to create reference years tied to certain percentiles, i.e., the median January “... combined with the median February, March, etc.”

One question that arises in the creation of any synthetic data is its advantages over recorded data. If long-term high-quality data is available for some location, is there

any point in using synthetic data? As Kershaw et al. (2011) point out, the utility of recent records in predicting future return periods (i.e., probabilities of weather events of interest) is limited by the length of the record. For example, if a 100-year event (over a long enough record, this event will occur roughly 1% of the time) happened thrice in the last 10 years, does that make it a 3-year event or not? While the return period obtained from any weather generator is speculative, it does at least provide bounds on a system’s response. Then, it is the decision-makers who must choose the probability for which they would like to design. For example, HVAC system failure may be acceptable for some value of outdoor temperature or episode of some intensity, which has a very low probability of occurrence. Kershaw et al. (ibid.) warn that using the UKCP09 weather generator to assign return periods should be done with “extreme care”, and “...return periods longer than 5-years should be used with caution”.

A long record does enable a sensitivity analysis, but one is still hostage to the vagaries of the weather when using it. That is to say that there are several possible future conditions that may not have occurred in the recent past. There are no guarantees about what conditions may prevail in the future based on knowledge of past conditions. As far as we are aware, the temperatures of future years do not have to follow some well-defined mathematical relationship with temperatures from previous years, or even some well-defined periodic relation. The intention of our work related to the creation of synthetic weather data is not to *predict* future weather. Incorporating stochasticity does not automatically improve the predictive power of simulation for a specific time in the future. Rather, we expand the role of simulation in exploring design options by broadening the test conditions.

This paper begins with an explanation of the method used to construct future weather time series. We then present some descriptive statistics about the generated series. Finally, the results of simulating a single family home with these ‘future time series’ are compared to simulation with recorded data from the last two decades. This building model has been previously described in Chinazzo (2014).

METHOD

The work presented in this paper builds on previous work by the authors (Rastogi and Andersen 2015). We begin with a brief overview of previous work, avoiding repetition as far as possible. The generation of synthetic future weather data presented here relies on two major steps: time series decomposition and resampling¹. Both steps are explained in detail in our previous work, and summarised here in fig. 2. This paper uses the same terminol-

¹We use the terms *resampling* and *bootstrapping* interchangeably in our work. See Davison and Hinkley (1997) and Politis (1998) for a refresher on the theory, or Rastogi (2016) for a summary.



Figure 1: The single family home simulated as an example, details of which are in Chinazzo (2014).

ogy as previous work, as does fig. 2.

Previous Work

Several publications, listed in Rastogi (2016) and Rastogi and Andersen (2015), have showed that temperature, solar radiation, and humidity can be divided into periodic and aperiodic components. That is, if the periodic part of the original time series is removed (by subtracting Fourier series with appropriate periods, μ_t and ζ_t , for example), the remainder is aperiodic *noise* (ϵ_t). These two parts are shown in fig. 3 for Geneva. This noise is not entirely free from structure, however, and is generally well described by low-order Seasonal Auto-Regressive Moving Average (SARMA) models ($\psi(L)$). Upon fitting these low-order models, the residual is near-white noise. This residual or remainder term (r_t) is reshuffled, in 3-day blocks separated by month, to create new ‘resampled’ residuals (\hat{r}_t). These resampled series are input as the noise component in simulating the fitted SARMA model to create new synthetic aperiodic noise components $\hat{\epsilon}_t$.

A Seasonal Auto-Regressive Moving Average (SARMA) model is a combination of seasonal and non-seasonal Auto-Regressive (AR) and Moving Average (MA) terms. The idea is that the value of a time series at a certain point in time is predicted by a polynomial composed of two parts: an AR part and an MA part. The AR part is a regression of a time series on itself, i.e., its own values in the past. The MA part is averaged white noise, a weighted average of a finite number of white noise draws preceding the current time step. The seasonal terms include every Q^{th} past value, e.g., every 24 hours back from the present. The non-seasonal terms refer to the last p terms, e.g., 1-4 hours ago (generally p for AR and q for MA). The coefficients of the polynomial are estimated using maximum likelihood estimation to minimise the residuals, $r_t \sim \mathcal{N}(0, \sigma)$. The details of these can be found in a time series analysis book like Cryer and Chan (2008),

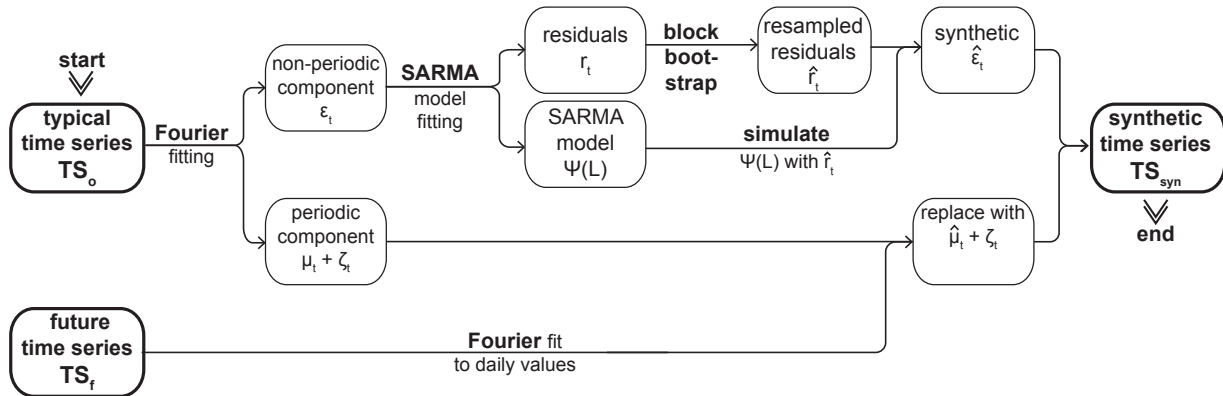


Figure 2: Generating synthetic weather time series from typical data and future forecasts of daily mean values.

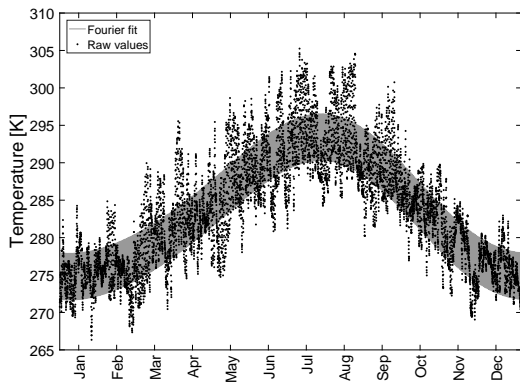


Figure 3: The periodic part of TDB series for Geneva (line) is overlaid on the raw hourly values (dots). Three pairs of Fourier terms are used to create the periodic signal: 8760 hours, 4380 hours, and 24 hours.

or the documentation of a software like the ones we used: MATLAB[®] `arima`, `estimate`, `infer`; and the forecast package in R (Rob J. Hyndman and R Core Team 2015).

Periodic Signal

Previous work to create future weather files focussed on a ‘fixed’ addition of some forecast to current data (morphing) or the creation of future data through relationships with a single forecast variable (e.g., rainfall for the UKCP09 projections). The most popular approach, morphing, is limited to producing “... a future weather pattern... that is largely analogous to the present-day weather in terms of diurnal cycles and extremes” (Jentsch et al. 2008). In our work, the combination of a random SARMA model and climate change forecasts creates ensembles of future time series, each of which is unique. As with any data-based methods, we cannot actually account for the changed physics of the atmosphere. That is what the Global Climate Model (GCM)-based forecasts are meant

to simulate. The stochastic models added on to the low-resolution future series create variation around this forecast, generating (bootstrapped) confidence intervals.

The periodic parts of the meteorological time series being considered, TDB and Relative Humidity (RH), are generally composed of a low-frequency signal and a high-frequency signal. The temperature series needs three Fourier pairs: one for annual seasonal variability, or a pair of terms with a period of 8760 hours; one for diurnal variability, or a pair of terms with a period of 24 hours; and, an additional pair with a period of half a year or 4380 hours, to shift the peak slightly to the right of centre towards August. Humidity shows no appreciable diurnal variations, so reduces to aperiodic noise with the removal of just an annual signal. We decided to not de-trend and simulate the global horizontal irradiation series separately since it creates additional artefacts that are difficult to remove without extensive, manual, post-processing.

The climate change forecasts available to us were daily mean values for this century (up to 2100). Two Representative Concentration Pathways (RCPs) were explored in our study, RCP 4.5 and RCP 8.5, details of which can be found in *Climate Change 2014*, pg. 8. The first corresponds to an intermediate emissions scenario while the latter to one with very high Green House Gas (GHG) emissions. These RCPs are simulated using GCMs, which are downscaled by meteorological agencies for their regions of interest. We had access to the Regional Climate Models (RCMs) for Europe through the CORDEX project website (World Climate Research Programme 2015). There are several GCM model runs available on the CORDEX website forecasting each emissions scenario for Europe, all of which can be considered equivalent. That is to say, there is no claim that any one model is more accurate or likely than another.

Incorporating Forecasts

The process begins with a selection of forecast daily values from one of the GCM/RCM model runs for either RCP 4.5 or RCP 8.5. One may pick any one of these model combinations to create a ‘string’ (or ensembles of strings) of 85 years (2015-2100), or use an average. For this study, only one GCM/RCM model combination of daily values is used for demonstration. Each string corresponds to either of the two RCPs, since each RCP separately represents a possible future outcome under a consistent set of assumptions.

These future daily values are used to replace the low-frequency Fourier series μ_t . Conceptually, a Fourier fit with a period of 365 days fit to daily values is identical to one with a period of 8760 hours fit to hourly values. Thus, we can insert this ‘future’ low-frequency signal instead of the ‘present’ low-frequency fit in the reassembly of a complete future time series. Referring back to fig. 2, instead of putting back the original μ_t to get the *plain* TS_{syn} , we put in a different $\hat{\mu}_t$ that represents future daily mean values. If present, the daily term ζ_t remains constant. Finally, adding the simulated noise values ($\hat{\epsilon}_t$) creates any number of variants (weather years) for a given combination of a future series ($\hat{\mu}_t$) and the (unchanged) daily signal ζ_t , for a particular RCP. This procedure is used for TDB and RH, while the solar terms, Global Horizontal Irradiation (GHI), Diffuse Horizontal Irradiation (DHI), and Direct Normal Irradiation (DNI), are created using a nearest-neighbour bootstrap described below.

Like we mentioned in previous work on creating ‘plain’ synthetic files (i.e., without climate change forecasts), these synthetic ‘future’ time series have to be cleaned (censored) due to the nature of the generation process. For example, the SARMA simulation added to the periodic signal may create a final value of 70°C for TDB or 0.5 for RH, because the procedure does not ‘know’ that these values are invalid. Even if these physically invalid values are removed, there are still ‘outliers’ seen in some series upon visual inspection. The definition of outliers is a complicated matter, so we use historical data as a guide. For example, if a certain hourly change in temperature is seen in the source Typical Meteorological Year (TMY) file, then we assume that it is *possible*. Meaning that the raw TDB values that caused this change need not be censored. There are several different techniques to remove outliers, of which we used a method based on standard z-scores,

$$z_i = \frac{x_i - \bar{x}}{s}, \quad (1)$$

where, \bar{x} is the sample mean and s is the sample standard deviation. By itself, this score does not indicate that a particular data point is an outlier. Rather, an arbitrary cut-off point must be decided. The advantage of using z-scores is

that instead of imposing arbitrary limits on the raw values of a parameter, which are highly climate or context dependent, it is possible to use standardised values and cut-offs. This helps to maintain consistency across climates and parameters. In our case, we found that choosing the larger of the 99.9 and 0.1 percentiles is sufficiently conservative to remove outrageous values (like 100°C or -100°C) but not so conservative as to remove extremes. The time series is censored for both high and low values. This is obviously an arbitrary choice, and the generation of weather files is only moderately affected if this cleaning is not carried out. We looked at various cut-off values, and could not arrive at a conclusively universal one. This is because we do not take a position on *which extreme is too extreme*. We expect that visual inspection or expert opinion is as good as hard-coded checks in the generator. Most building simulation programs have their own cut-offs for valid values, but since we use daily mean TDB in subsequent steps, censored values are easier to work with. We censor both raw hourly values and the first difference of the series (hourly changes in values).

GHI, DNI, and DHI are treated differently from the others since no attempt was made to fit and remove a periodic component from any solar time series. As these three series are dealt with in tandem, we will discuss only the production of a synthetic time series for GHI. Instead of fitting models to create synthetic hourly values for solar radiation, we decided to resample from the values available in the TMY file. There is a reasonably strong correlation between the daily sum of GHI and daily mean of TDB, as evidenced by values of 0.7-0.75 for Pearson’s (linear) correlation coefficient (r) and Spearman’s rank correlation coefficient (ρ) in most climates. This should not be over-interpreted to mean that daily mean TDB is necessarily well-described by a linear function of the daily sum of GHI. Rather, it is an indication that, in addition to the effect of the season (which is an indication of the ‘band’ of temperatures within which most values in a month will lie, and the hours of radiation in a day), high solar irradiation during the day will generally coincide with higher mean temperatures. In this case, we are merely exploiting this correlation to find valid day-long series of solar radiation. Belcher et al. (2005) point out that there do not seem to be any mechanisms in climate change models causing massive shifts in the amount of solar radiation delivered day by day. What might change for some climates is the number of cloudy or partially cloudy days, leading to a change in the quantum of solar radiation received over a long enough period like a year. Since the length of day and maximum values of GHI are related to latitude, altitude, and the solar constant, none of which are affected by atmospheric concentration of GHGs the authors propose that it is valid to use past or typical data as the source of

future data.

The process of selecting future ‘solar days’ is split by month, since the length of the day depends on the time of year. For each day in a given ‘future’ month, we calculate the daily mean TDB. Then, we locate k days in the TMY file, in the same month, that have the closest daily mean temperatures to the future daily mean temperature being considered. Of these k days, any random one is chosen for its solar profile (i.e., hourly solar data), which becomes the hourly data for the future day. In this way, hourly temperature values (represented by their daily means) that have already occurred with certain hourly solar values, occur in the future files as well. Some noise has been introduced in the process by initially calculating the k nearest neighbouring days to a future day, in the same month, and then choosing one randomly. In our work, we used $k = 10$ nearest neighbours. While the forecasts available to us do include future daily mean GHI values, we chose not to work with these since their distribution was not very different from the daily mean GHI values seen in the source TMY file. Basing the selection of future hourly values purely on mean GHI forecasts (i.e., by locating the same value in the TMY file) could break the cross-correlation between TDB and the solar time series.

Variants

Our method creates any number of variants for a given year by simulating the SARMA model with bootstrapped residuals, as described in our previous work (Rastogi and Andersen 2015). The final values are a combination of the synthetic residuals $\hat{\epsilon}_t$, unchanged daily Fourier term ζ_t , and future forecasts $\hat{\mu}_t$, which are equivalent to the low-frequency Fourier term μ_t . Thus, each ‘string’ of 85 years is based on one GCM/RCM model and a set of simulated residuals. The individual variants for each year, of which there can be any number, are an ensemble representing the possible weather that may occur in the future. They are meant to be used together, not individually, since the authors do not claim that any one variant is more likely than another. The nature of Monte Carlo simulation dictates that a small sample size, or even worse a single sample, is almost certainly not representative of the phenomena being simulated. Similarly, the future time series should also be interpreted loosely – neither climate change forecasts nor our methods are precise enough to predict a specific value in the future. A simple rule of thumb we propose is that the years of each decade should be treated as being interchangeable.

RESULTS & DISCUSSION

Raw Weather

Figure 4 shows the empirical Cumulative Distribution Functions (eCDFs) of synthetic TDB and RH, alongside recorded values and TMY. The distributions are virtually

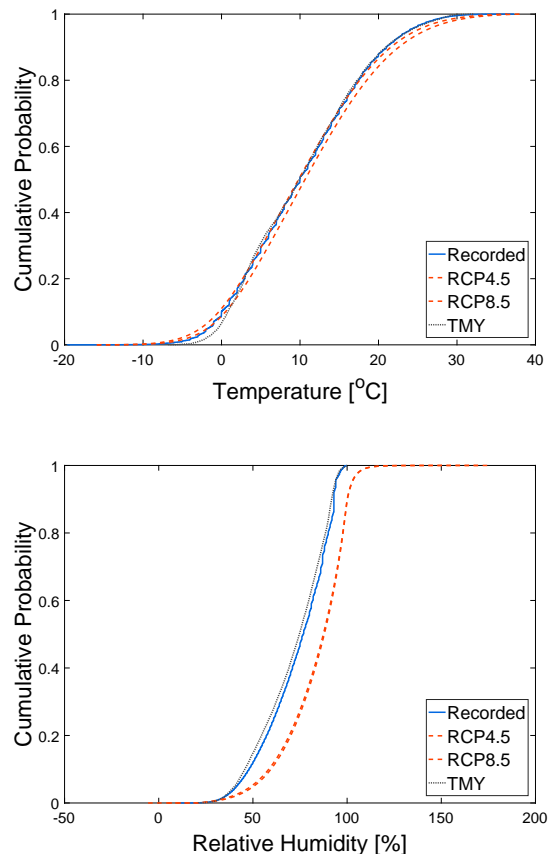


Figure 4: The eCDF of hourly TDB [top] and RH [bottom] values for TMY, recorded, and future data.

identical. The monthly extents of values seen in the synthetic and recorded data are given in fig. 6. These figures, and the percentiles given in table 1, show that the extreme temperatures and humidity values seen in the past 30-odd years of recorded data are well reproduced (and exceeded) in the synthetic series.

Looking at the monthly extents (table 1 and fig. 6), it seems that both the recorded data (1984-2014) and climate change forecast-based files are slightly warmer than the design temperatures, i.e., the lower extremes are less extreme and the higher extremes are worse. The fact that the recorded data contains warmer extremes is no surprise when considering that the TMY file for Geneva is composed of months from the 1980s and 1990s, whereas the 2000s have broken several high temperature records. Depending on the source of the typical weather files for different locations, the ‘baseline’ or source years may be even older. However, the absence of winter extremes should not be taken as a given: climate change forecasts do not simply ‘shift’ the existing weather data upwards. The occurrence and intensity of extreme events, e.g., very low or very high temperatures, is unpredictable with anthropogenic climate change, since past records are less

Table 1: ASHRAE design temperature percentiles for Geneva. All TMY values are taken from the header of the TMY file, except for the 98th percentile. This was calculated, and so represents the 98th percentile of the ‘mean’ signal.

Perc. (%)	Geneva – 50-sample run				
	Rec.	TMY	Syn.	RCP4.5	RCP8.5
99.6	31.13	30.05	30.80	32.56	34.43
99.0	29.21	28.33	29.00	30.24	31.93
98.0	27.35	26.80	27.20	28.00	29.56
50.0	10.10	10.00	10.41	9.66	10.53
2.0	-2.77	-3.70	-1.90	-4.85	-4.09
1.0	-4.04	-5.00	-4.80	-6.53	-5.80
0.4	-5.63	-7.20	-6.90	-8.56	-7.82

representative of the future. On the other hand, the ‘average’ temperature is marching inexorably upwards.

The broad agreement between new data and our synthetic climate files would suggest that the synthetic approach is usable for simulating diverse future conditions. The disagreement of the synthetic data with TMY data is also a plus, as explained above. We should point out, however, that the ‘plain’ synthetic files (i.e., ones that did not incorporate forecasts) also did a good job of producing extreme values (as reported in our previous publication). Our preliminary conclusion is that the appearance of extremes depends more strongly on the SARMA simulation, so it occurs with or without the inclusion of climate change forecasts.

Simulation Results

The four different kinds of weather input files shown in fig. 7 are: recorded files, which include typical files from the United States Department of Energy (USDOE) website and the METEONORM (MN) software; plain synthetic files, which do not include climate change forecasts; and synthetic files incorporating projections from the two RCPs under consideration. The simulation results show very similar distributions. That is to say that the extents of the spread, and its shape, are roughly equal for the various kinds of files. The RCP4.5 files have a more skinny distribution because fewer of them were simulated. The RCP8.5 values show the largest extents. In general, the synthetic files (both plain and with forecasts) show extremes comparable to or bigger than the recorded data. This is an important result: the synthetic files reproduce extremes *near* the ones seen in the past 30-odd years, with nearly the same probability, and extend them a little further. We are confident that larger samples of synthetic files (plain and future) will show extremes of longer return periods (i.e. lower probability).

We expect that the annual sum of heating or cooling en-

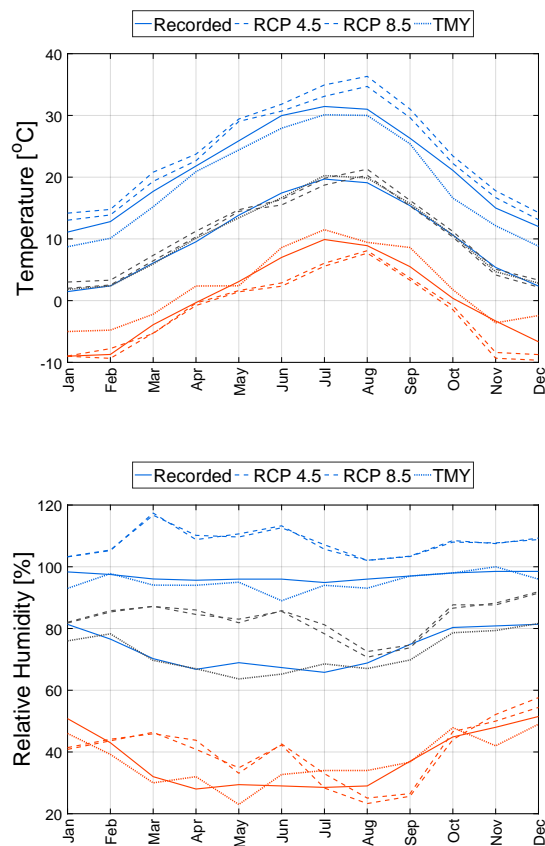


Figure 6: The extents of hourly TDB [top] and RH [bottom] values, by month, for TMY (dotted line), recorded (solid), and future (dashes) data. The upper lines represent monthly maximums (99th percentile), while the lower lines are for monthly minimums (1st percentile). The lines in the middle are for monthly means. The synthetic data extremes are appreciably higher, but the probability of those extremes is still as low as in the recorded data.

ergy usage should be more sensitive to shifts in overall temperatures rather than the occurrence of intense events, so the significant overlap between plain and future synthetic files is somewhat surprising. We expected that the addition of an upward signal would change the overall energy usage appreciably. Looking at fig. 8, we see the reason that this is not apparent in fig. 7. The range of values possible in the future, i.e., the spread due to different weather possibilities in the same climate, is so large as to drown out the gradual shift seen year-by-year. So, while the prediction is for a gradual warming of the climate, the uncertainty in future values makes forecasting noticeable reductions in heating (or increases in cooling) very inaccurate. In upcoming work, the authors are analysing other metrics such as peak demand and overheating to assess if useful predictions can be found for those. For example,

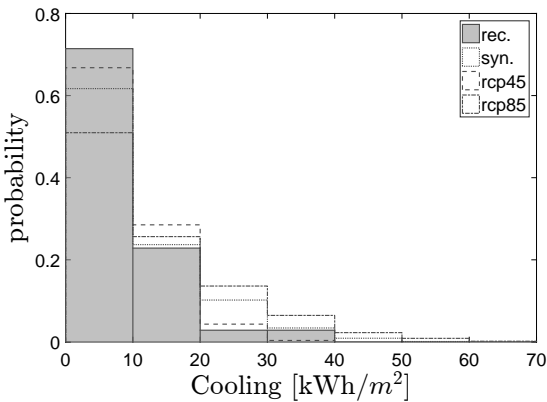
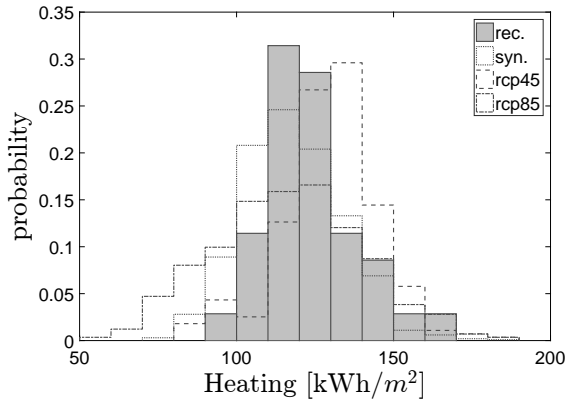


Figure 7: Histograms for EUI heating [top] and cooling [bottom]. The distributions of the four different kinds of weather files are nearly identical.

the frequency of future extreme events described in Ker-shaw et al. (2011).

CONCLUSION

In this paper, we have explained our method for incorporating climate change forecasts into an overall schema for generating synthetic weather files. The use of these files is primarily to enable the exploration of what-if scenarios, vis-à-vis weather, to get a range of possible outcomes (e.g., range of annual cooling energy used). So, while it is instructive to compare the synthetic data to recent recorded data, the generation process is meant to also create values that have not been seen before. The point of this exercise is not to predict the weather at a given point of time in the future, since that is beyond the ken of contemporary climate models. Instead, we are looking to provide a sufficient variety of physically-valid weather conditions based on GCM model outputs. Upon simulation, these conditions generate a statistically valid sample of outcomes, like energy use, to have an idea of the robustness of a building or design, an idea previously developed

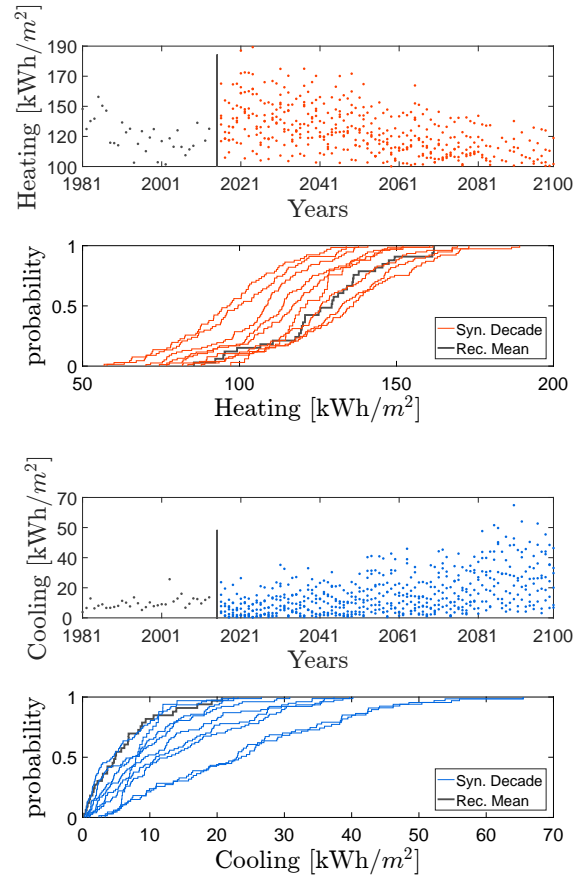


Figure 8: The EUI [kWh/m²] plotted by year, [top] heating, [bottom] cooling. The line at 2015 represents the extent of results from plain synthetic files. Cumulative distributions of annual energy use values in the next few decades, 2010-2100, are plotted below each yearly plot.

by the authors in Chinazzo et al. (2015a,b).

This paper demonstrates a method to include climate change forecasts with variation in individual values, but it cannot account for the physical effects of the build-up of GHGs in the atmosphere. Users must rely on climate models for that. For Geneva, the forecasts show a very small upward trend of temperature. The synthetic time series created using TMY files from the 1980s-90s and the newest climate change forecasts tally well with recent recorded data, which includes the 2000s. That is, the effect of including climate change forecasts on older data (the TMY files) is similar to actual recently recorded data. This was to be expected since the concentration of atmospheric GHGs has been increasing steadily for more than a century, the effects of which have only become apparent in the past couple of decades.

We have also discussed why our proposal is distinct from previous efforts based on morphing and similar techniques. While morphing is unable to produce files with

sufficient variety, we are able to produce very widely varying samples of weather from a future climate scenario rapidly. Like morphing and any other synthetic weather generator, it should be noted that our synthetic weather files are not explicitly accounting for geographical variability. That is to say, if a source TMY file is not representative of the building site (e.g., due to urbanisation), then our method will not correct for it. This is an important limitation, and one we will address only in upcoming work, since the ‘change’ in weather conditions due to urbanisation has nothing to do with the techniques we use here. It is possible to coincidentally reproduce urban conditions, but that is not guaranteed.

The method shown here, and in Rastogi (2016) and Rastogi and Andersen (2015), is also applicable when a long record of weather data is available. We have focussed on working with typical year files to expand applicability to practice. Longer, high-quality, records, where available, could be a better basis for calculating the various periodic and aperiodic components we use in our method. The influence of the quality of typical files is not formally addressed in our work, but the use of an ensemble of random files could ameliorate somewhat the impact of unrepresentative data on decision-making.

ACKNOWLEDGEMENTS

This work was carried out at the EPFL, and supported by the CCEM-SECURE project and the EuroTech Universities Alliance. The advice of Prof. A.C. Davison and M. Kuusela has been invaluable in the development of this work. G. Mavromatidis’ help in obtaining and interpreting the climate change forecasts is gratefully acknowledged, along with his support. The large number of simulations shown here would not have been possible without the help of Dr. R. Evins.

References

- Belcher, SE, JN Hacker, and DS Powell (2005). “Constructing design weather data for future climates”. In: *Building Services Engineering Research and Technology* 26.1, pp. 49–61.
- Chinazzo, G (2014). “Refurbishment of Existing Envelopes in Residential Buildings: assessing robust solutions for future climate change”. MSc. Lausanne, Switzerland: EPFL.
- Chinazzo, G, P Rastogi, and M Andersen (2015a). “Assessing robustness regarding weather uncertainties for energy-efficiency-driven building refurbishments”. In: *Proceedings of IBPC 2015*. 6th International Building Physics Conference. Torino.
- (2015b). “Robustness Assessment Methodology for the Evaluation of Building Performance With a View to Climate Uncertainties”. In: *Proceedings of BS 2015*. 14th International Conference of the International Building Performance Simulation Association. Hyderabad, India.
- Crawley, DB (2008). “Estimating the impacts of climate change and urbanization on building performance”. In: *Journal of Building Performance Simulation* 1.2, pp. 91–115.
- Cryer, JD and KS Chan (2008). *Time Series Analysis: With Applications in R*. Springer. 501 pp.
- Davison, AC and DV Hinkley (1997). *Bootstrap Methods and their Application*. 1st. Cambridge University Press. 594 pp.
- Eames, M, T Kershaw, and D Coley (2011). “On the creation of future probabilistic design weather years from UKCP09”. In: *Building Services Engineering Research and Technology* 32.2, pp. 127–142.
- IPCC (2014). *Climate Change 2014 Synthesis Report: Summary for Policymakers*. Geneva, Switzerland: Intergovernmental Panel on Climate Change (IPCC).
- Jentsch, MF, AS Bahaj, and PAB James (2008). “Climate change future proofing of buildings - Generation and assessment of building simulation weather files”. In: *Energy and Buildings* 40.12, pp. 2148–2168.
- Kershaw, T, M Eames, and D Coley (2011). “Assessing the risk of climate change for buildings: A comparison between multi-year and probabilistic reference year simulations”. In: *Building and Environment* 46.6, pp. 1303–1308.
- Politis, D (1998). “Computer-intensive methods in statistical analysis”. In: *IEEE Signal Processing Magazine* 15.1, pp. 39–55.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rastogi, P (2016). “On the sensitivity of buildings to climate: the interaction of weather and building envelopes in determining future building energy consumption (in preparation)”. PhD thesis. Lausanne, Switzerland: Ecole polytechnique fédérale de Lausanne.
- Rastogi, P and M Andersen (2015). “Embedding Stochasticity in Building Simulation Through Synthetic Weather Files”. In: *Proceedings of BS 2015*. 14th International Conference of the International Building Performance Simulation Association. Hyderabad, India.
- Wilde, P de, Y Rafiq, and M Beck (2008). “Uncertainties in predicting the impact of climate change on thermal performance of domestic buildings in the UK”. In: *Building Service Engineering Research and Technology* 29.1, pp. 7–26.
- World Climate Research Programme (2015). *CORDEX*. WCRP CORDEX. URL: <http://www.cordex.org/>.