

EMBEDDING STOCHASTICITY IN BUILDING SIMULATION THROUGH SYNTHETIC WEATHER FILES

Parag Rastogi¹ and Marilyne Andersen¹

¹Interdisciplinary Laboratory of Performance-Integrated Design (LIPID),
Ecole Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland

ABSTRACT

This paper presents an attempt to create synthetic weather data for stochastic building simulation. The synthetic data are created entirely from the freely available Typical Meteorological Year (TMY) weather files using time series models and resampling. The generated data turn out to be representative of recorded data for our case study without any prior ‘knowledge’ of the long term distributions of meteorological parameters. The current model does not address spells above or below some temperature of interest (e.g. heat waves), and the authors are working to incorporate that in future work. Another avenue for further exploration is modifying the mean to incorporate the results of Regional Climate Models for future conditions. Correlation of the synthetic data with synthetic solar radiation and humidity has been verified and the authors’ work with this ensemble of weather time series of interest will be presented in future publications.

INTRODUCTION

Using TMY files in building simulation for ‘what-if’ analyses of designs tells us only about the response of a building or design strategy to typical climatic conditions, for the period of record of the file. However, several studies have pointed out the sensitivity of simulation output to weather data, including Bhandari et al. (2012), Chinazzo (2014), Crawley and Huang (1997) and Hong et al. (2013). In other words, simulating with typical weather gives no information about the sensitivity of a building or design strategies to variations in the climate itself.

The difficulty of fully characterising a system (for sensitivity or uncertainty analyses) that depends on the climate is that we cannot fully characterise the climate itself, especially future climate. This epistemic uncertainty has led some to propose that a ‘range’ of possible performance outcomes, i.e., the results from simulation runs with different weather inputs, better characterise the *range* of performance that a building will inevitably give (e.g., Chinazzo et al. 2015). If one does not know exactly what (weather) inputs one’s (building) system will experience, then one is better off knowing the effect on it of a range of possible inputs. A given weather file is, after all, a representation of one scenario out of an immense number of possibilities. Therefore, by using only one weather file, we

are restricting ourselves unnecessarily to one “experimental result”. If a building never experiences a narrow set of weather conditions exactly, i.e., the ones contained in a typical weather file, then the quality or ‘averageness’ (or ability to represent best the most typical weather) of said weather file is irrelevant.

The aforementioned studies usually propose using measured weather data from the vicinity of the building to best characterise the climate. While long simulation with measured data is, intuitively, a better depiction of climatic variability than a single ‘typical’ time series, it can not guarantee coverage of future conditions or extremes. While the historical distribution of a weather parameter can be known from recorded data (usually at a resolution of one day or month), using historical records to study the future represents an assumption about the future: that of a stable climate. The IPCC’s latest Synthesis Report (IPCC Core Writing Team 2014) is unambiguous in its assertion that the climate is changing, though it is not knowable which of its future scenarios best represents how the global climate will evolve.

Assuming one has access to long-term hourly data from a weather station that is sufficiently close to the area of interest, in addition to a TMY-type file, one can know how a building *would have* behaved. However, one has no tools for assessing any arbitrary weather conditions. One might have a reasonable idea of the expected range of average temperature rise in a climatic region, thanks to the IPCC’s publicly available models, but one does not know the possible implications of this at an hourly resolution for a given weather station. Our approach seeks to address this incertitude by proposing a ‘what-if’ analysis of a building to variations in the climate, without seeking to forecast the ‘true’ future climate.

The first step in the development of this climate sensitivity analysis procedure is introduced in this paper: stochastically-generated synthetic input data. We extract the essential characteristics of a climate (e.g., autocorrelation, means, etc.) and build any number of synthetic files by modelling the structures and reshuffling the apparently random components of the time series. This is possible because of the idea, developed by several authors (including, Boland 1995; Boland et al. 2013; Hansen and Driscoll 1977; Magnano et al. 2008; Scartezzini et al. 1990), that weather time series

can be decomposed into characteristic seasonal components and apparent “innovations”, though without any claim to know the source of these random changes or innovations.

Bootstrapping, which is what we use to create variations of the innovations, is a resampling technique. Resampling refers to any of a set of methods which compute the bias of an estimated (sample) statistical measure like the mean (Dodge 2008). These methods do not rely on knowing particular parameters describing the shape or distribution of the population or sample (e.g., mean, standard deviation, etc.). They also avoid the restrictions imposed on a parametric random model by a small sample size (the TMY in our case) and long sample runs. Resampling methods are more useful in non-parametric situations where a model is not available, and the data has to “do all the talking”. Essentially, “the bootstrap method amounts to treating your observed sample as if it exactly represented the whole population” (Politis 1998). On the other hand, resampling methods are harder to use when dealing with time series that show a high degree of seasonality or correlation. In other words, all deviations from an *iid* (independent identically distributed) assumption must be dealt with. For example, temperature on a July night is highly correlated to a July day, but not to a January night. A resampling run cannot distinguish between day and night temperatures or summer and winter – a fatal problem for weather data. Since resampling by itself does not maintain cross- and auto-correlations, we propose ways to overcome this in our work.

In this paper we show examples of synthetic time series generated from temperature values from a TMY file. The synthetic series are compared with actual weather data using a number of tests. In this paper, we do not foray into simulating a change in mean monthly temperatures. That work is in progress, and will be presented in future publications. Similar treatment of Global Horizontal Irradiation (GHI) and Humidity Ratio (W) is not presented in this paper due to lack of space. The authors do not plan to assess the suitability of this treatment to wind data. There are two reasons for this: wind data is less relevant in the context of building simulation unless natural ventilation is to be assessed, and the wind data in a TMY file cannot be considered to be representative since the wind time series is not used in the selection of typical months. Studies simulating wind velocity for power generation include Olsina (2013), who created synthetic data based on the spectral representation of a short record.

The implications of the number of synthetic files generated (n_{sim} or n_{boot}) is not explored in this paper, and we worked with 100 realisations for this demonstration. A quick analysis of the effect of increasing the number of simulation runs up to 1000 showed that there is no appreciable improvement in the performance of the modelling procedure with more realisa-

tions. Future work will look into finding the minimum possible realisations that generate some weather conditions of interest (e.g., a summer extreme).

METHODOLOGY

In this paper, we demonstrate our method with a Typical Meteorological Year (TMY) file from Geneva, Switzerland. We are only concerned with three time series from these files: Dry Bulb Temperature (TDB), GHI, and Relative Humidity (RH). However, there is no reason to doubt its applicability to other parameters like Direct Normal Radiation and Cloudiness. However, the cross-correlation or dependence structures become progressively complicated with several series. The method discussed in this paper is a significant improvement from the exploratory method and data reported in the authors’ previous work (Rastogi and Andersen 2013). The method also borrows heavily from the work of John Boland (Boland 1984; Boland 1995) and Magnano et al. (2008). In particular, a significant insight proposed in their latest paper greatly improves model fit: that of using two levels of models, a high-frequency fluctuation and a low-frequency one. The major difference between the models proposed by these authors and our model is twofold: we restrict ourselves to using TMY or Design Reference Year (DRY) files, while they used recorded data; and we aim to demonstrate a generally applicable model for creating synthetic weather data for simulation, whereas they were only working with a sample of stations in Australia. Other differences between our method and theirs are mentioned in the text below.

All statistical procedures mentioned in this paper were carried out using MATLAB[®], with a brief excursion into R[®]. The descriptions of some procedures, statistical tests, etc. in the text are based on the documentation of MATLAB[®]. Much of the work is also based on the discussions with Prof. Davison (Davison 2003). We begin with an explanation of the hourly temperature model and the method of generating synthetic time series. Then, we test the output data against TMY and recorded data for Geneva.

Hourly Temperature Model

We begin by converting the TDB values to Kelvin, which makes the whole series strictly positive. This greatly simplifies error correction and post-processing of the data at the end. Our hourly model, given by (1), is similar to that of Magnano et al. (2008, eq. 3).

$$x_t = \mu_t + \zeta_t + \epsilon_t \quad (1)$$

One difference is in the first term: while their \bar{x}_j represents the *calculated* daily means from recorded hourly temperature values, our μ_t is a low-frequency one-term Fourier fit representative of the seasonal fluctuation. The subscript t is hour of the year, ζ_t describes the high-frequency daily cycle and the residual, ϵ_t , is assumed to include some structure (e.g., autocorrela-

tion). The terms μ_t and ζ_t together constitute the ‘deterministic’ part of the total signal. Boland (1995) interprets the μ_t as the “long-term serial correlations” of temperature, or the effect of seasons. Magnano et al. (2008) suggest fitting a two-term Fourier model to the ζ_t term. They do this for each month, and for three bins per month. The bins separate the hourly temperature into: “moderately cold days (MC) when daily mean temperature is below or equal to 20°C; mild days (MD) when the daily mean temperature is greater than 20°C and less than or equal to 25°C; and hot days (HD) when the daily mean temperature is greater than 25°C”. The logic behind this is sound: the variance has an annual seasonal trend, and the daily magnitudes have a rough positive correlation with the daily mean temperature. Or, “the differences between the maximum and minimum are larger on days with higher daily mean temperature”. Another way to look at this is that the magnitudes of the deviations from the seasonal mean temperature values are not necessarily the same in each season despite removing the daily averages (which are in themselves characteristic of a season). For Geneva, the deviations do not follow an obvious pattern. Attempts to obtain a clear correlation of this estimated standard deviation with mean daily temperature or daily sum of solar radiation (a proxy for cloudiness) were unsuccessful. The daily sinusoidal nature of temperature is not evident in those 24-hour periods where clouds severely impact the radiant heat exchange of the earth with the sky.

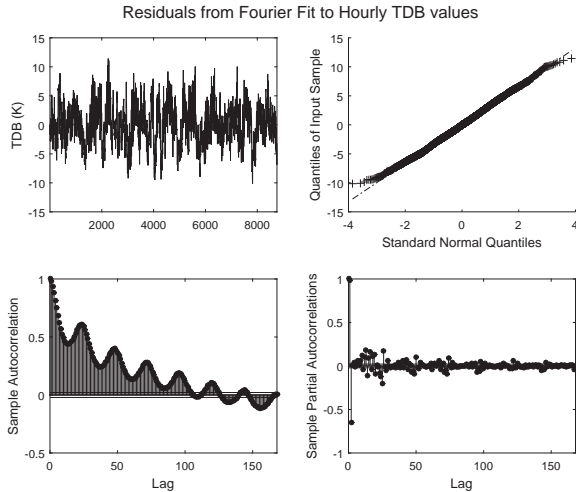


Figure 1: Descriptive plots for residuals ϵ_t from the Fourier fit to hourly temperatures. [Top left] Raw residuals; [Top right] Quantile-Quantile plot to check normality; [Bottom left and right] plots of Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF), i.e., correlograms.

We de-construct the hourly model sequentially: starting by subtracting the low- and high-frequency seasonal components from the hourly temperature values using Fourier series. The results of fitting a three-term Fourier model to the data are shown in Fig. 1. The frequencies of the fit were fixed at $2\pi/8760$, $2\pi/4380$,

and $2\pi/24$. The first and last frequency make sense physically. The last is less intuitive, but it helped to make the Fourier representation asymmetric by shifting peak summer a little after July, which gave a better fit. The ACF and PACF plots of Fig. 1 show residual Auto-Regressive (AR) and Moving Average (MA) structures. The periodogram (Fig. 2) of the residual shows leakage about the aforementioned frequencies. The upshot of these diagnostic plots is that the residuals can not be considered to be white noise.

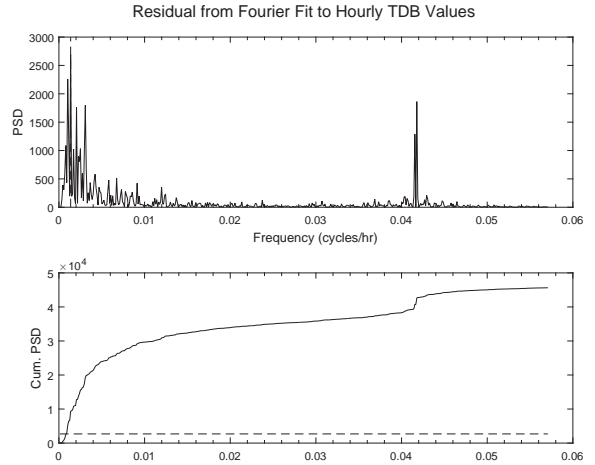


Figure 2: Raw and cumulative Power Spectral Density (PSD) of the residuals from the Fourier fit. High PSD values around $\omega = 0.0417$, or approx. 24 hours, correspond to leakage about the daily fluctuation. Leakage on the left of the plot corresponds to unknown waves of sub-yearly length.

We fit a $SARIMA(p, d, q) \times (P, D, Q)_s$ to remove the final remnants of structure from the ϵ_t term in (1). The Akaike Information Criteria (AIC) /Bayesian Information Criteria (BIC) scores, Log-Likelihood values, and characteristics of the residuals served to pick an appropriate model. We varied the AR, MA, Seasonal Auto-Regressive (SAR), and Seasonal Moving Average (SMA) lags between 0 and 4, with and without a seasonal integration factor of 24 (i.e., one day). Note that the term **seasonal** here refers to the daily cycle, not the seasons of a climate. The following trends were noteworthy:

- Models with a differencing factor of 1 day ($s = 24$) do not perform better than those without.
- Models with SMA and SAR lags of 24 hours do better than those without ($P = Q = 24$).
- Models with non-seasonal AR and MA lags ($p \geq 0, q \geq 0$) performed better than those without either, or without both.
- The best fitting models still fail to remove significant structure at lags of less than 72 hours (3 days). This is visible from the weakly significant ACF and PACF coefficients in Fig. 3

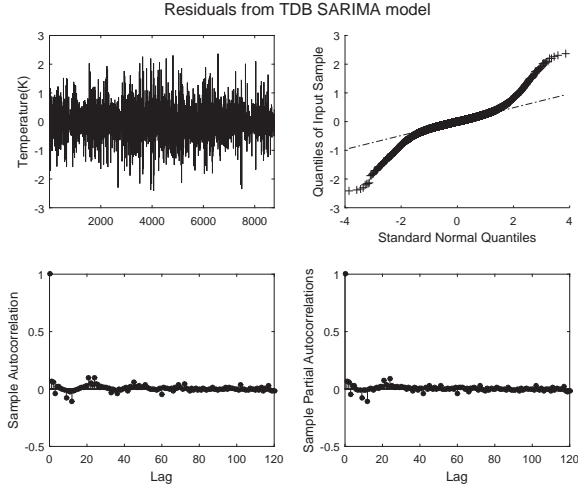


Figure 3: Residuals (ϵ_t) from fitting a model with Gaussian innovations to the ‘de-seasonalised’ hourly TDB time series for Geneva.

The best performing model for Geneva is $SARIMA(4, 0, 2) \times (1, 0, 1)_{24}$ with constant variance, i.e., a Seasonal Auto-Regressive Moving Average model (SARMA) model. The SAR and SMA lags are reported as multiples of their common period s (i.e., $SMA = 24$ is $Q=1$) using the Box-Jenkins convention. The residuals of this fitted model (Fig. 3) are treated as estimates of the innovations (or inexplicable randomness) in the ϵ_t term in (1). Thus the ϵ_t is broken up into an ARIMA polynomial, $\psi(L)$, and innovations, r_t , as shown in (2). The μ term represents a mean in formal notation, although in this case the time series has a zero mean. Like at various points in this procedure, the residual terms were also censored using the Z-score, with a cut-off of 10.

$$\epsilon_t = \mu + \psi(L)r_t \quad (2)$$

It is not possible to bootstrap individual values of the time series at this point to obtain samples of the r_t term from (2) because the structure at lags smaller than 72 hours is still present in the candidate series (Fig. 3). It is necessary to maintain this ACF not only to remain faithful to the original series but also to return a physically valid time series. Using a ‘block bootstrap’ instead of reshuffling individual points has a better chance of preserving short-term auto-correlations (Davison and Hinkley 1997; Politis 1998). The bootstrap is done using 3-day blocks, which is suggested by Magnano et al. (2008). Another restriction we imposed on the bootstrap was to subdivide the blocks by month and limit the reshuffling to blocks within a month. For example, January blocks were resampled with other January blocks but not those of July. This helped to preserve any residual sub-yearly correlation. We chose a 3-day block after exploring a few different options (1-7 days) to see which block length reproduces the auto-correlograms and partial auto-correlograms best. The different block lengths do not show substantially different results for the compar-

isons we use. Unsurprisingly, any block length preserves the intra-block ACF, but the behaviour up to 7 days is not consistent: the original TMY series decays faster than the synthetic ones. The PACF presents a slightly different story, in that the significant lags at 24, 48, and 72 hours are mildly exaggerated in the simulated series. The differences between the correlograms are visible but quite small. We are working to improve this, especially by checking the interaction of the bootstrap with the SARMA simulation.

The synthetic hourly temperatures are given by (3), where \hat{x}_t represents the synthetic values and $\hat{\epsilon}_t$ represents the simulated/bootstrapped final residuals. The $\hat{\epsilon}_t$ term is composed by simulating the SARMA model \hat{s}_t with the bootstrapped residuals \hat{r}_t . Since the bootstrapped residuals serve as the noise/innovation input to the simulation of the model, there is no additional noise introduced in the simulation.

$$\hat{x}_t = \mu_t + \zeta_t + \hat{\epsilon}_t \quad (3)$$

Generating Synthetic Series

The seasonal SARMA model was simulated using the outcome of a bootstrap run as the noise input each time ($n_{sim}=n_{boot}=100$). Some censoring of these simulated values was necessary! However, the 99th and 1st percentiles of the simulated values (\hat{s}_t) tend to be close to the maximum and minimum of the values in the original series (s_t). We removed the outliers using a modified Z-score proposed by Iglewicz and Hoaglin (as cited in NIST 2013). This handbook recommends treating any values which have an absolute Z-score greater than 3.5 as potential outliers. This was far too conservative for our purposes, so we ended up using the 99th and 1st percentiles as cut-offs. The Z-score is given by:

$$Z_i = \frac{0.6745(a_i - \tilde{a})}{MAD}, \quad (4)$$

where the Median Absolute Deviation (MAD) is $median(|(X_i - median_j(X_j))|)$.

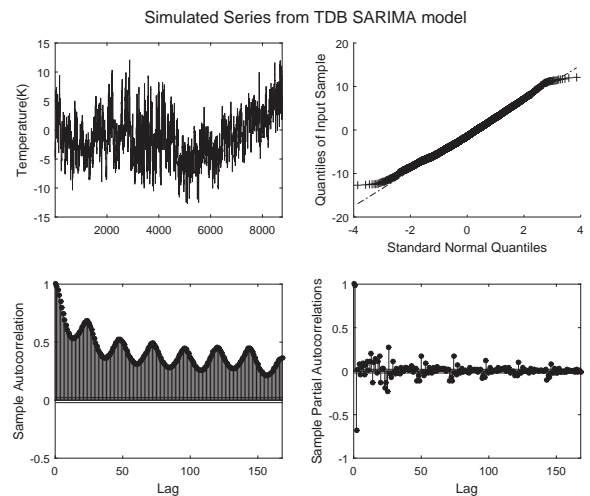


Figure 4: A randomly selected series of simulated \hat{s}_t values from the SARMA model.

The plots of Fig. 4 show acceptable similarity to the original hourly values of Fig. 1. The ACF decays much slower than the original but the ‘humps’ at the first three multiples of 24 hours are reproduced well. The resampled series (a random example is given in Fig. 5) show a remarkable similarity to the original sample, i.e., the residuals from fitting a conditional mean model to the ‘de-seasonalised’ series (Fig. 3).

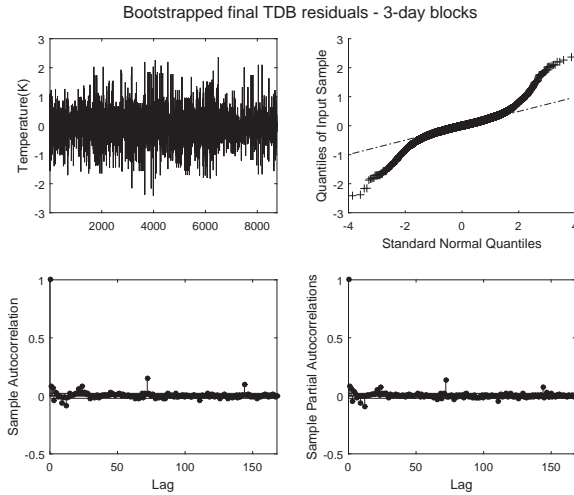


Figure 5: A randomly selected series of residuals (\hat{r}_t) obtained from bootstrapping the r_t term of (2).

Post-Processing

Some post-processing of the data is inevitable. Magnano et al. (2008) smooth the edges of the bootstrapped blocks to ensure that the edges of the blocks do not have an unacceptably large difference between consecutive values. We used the maximum/minimum first difference seen in the recorded data as the limits of what is an ‘acceptable’ hourly change of temperature. This gave us time steps we could classify as outliers, and the actual temperature values corresponding to these outliers were replaced with linearly interpolated values. We also censored the final synthetic time series using a conservative Z-score cut-off of 3.5, which is the value recommended in NIST (2013).

RESULTS

Table 1 shows the values corresponding to the various American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) design temperatures. These values, taken from the header of the TMY file, correspond fairly well to the percentiles calculated from the recorded data. Except for the 98th percentile, which was calculated from the TMY file itself, and so represents the 98th percentile of the ‘mean’ signal.

Table 1: Values corresponding to the ASHRAE design temperatures. The 99% design temp. is the 1th percentile, and so on.

Percentile	Synthetic	Recorded	TMY
99.6	32.78	30.05	30.80
99.0	30.33	28.33	29.00
98.0	28.21	26.80	27.20
50.0	10.07	10.00	10.41
2.0	-2.67	-3.70	-1.90
1.0	-3.90	-5.00	-4.80
0.4	-5.37	-7.20	-6.90

We now proceed to assess the results of the synthetic weather data generation using criteria from Boland (1995), Hansen and Driscoll (1977), Lund (1995) and Magnano et al. (2008). The comparisons were all done using recorded, TMY, and synthetic data. The measured data used in all the comparisons is from the period 1955-2014 for Geneva (MeteoSwiss 2014; NCDC/NOAA 2014), though with significant gaps. The gaps do not appear to have a bias. For example, there is no one particular season that is consistently missing. Pending further investigation, we assume that the measurement errors are uniformly distributed and do not colour the statistical characteristics of the time series. When the data gaps were too large (e.g., a whole month in a year), we removed that entire year.

Measures of central tendency and dispersion

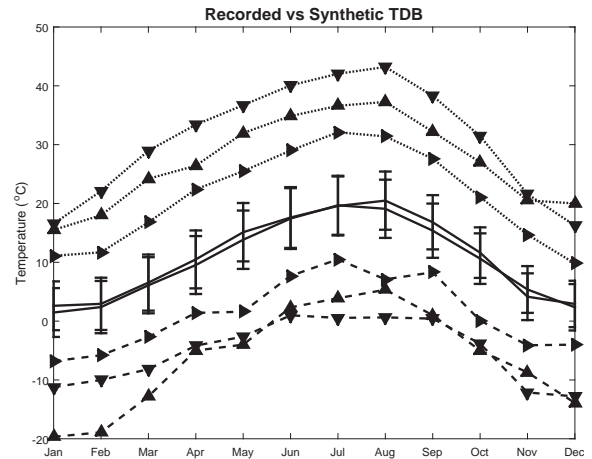


Figure 6: Summary statistics: recorded data are represented by upward-facing triangles, TMY by right-facing, and synthetic by downward-facing. The maximum temperatures for each month are plotted with a dotted line, the means with a solid line, and the minimums with a dashed line. The bars around the mean temperatures represent $\pm\sigma$, or one standard deviation.

Fig. 6 shows a comparison in terms of range (maximum, minimum) and the first and second moments (mean and standard deviation). Fig. 7 shows the spread of the data (median and quantiles). In general, the approach characterises measures of central tendency (mean and median) and some dispersion (standard deviation and inter-quartile range) very well, but is not

able to reproduce extreme values (see the difference in the lowermost lines of Fig. 6). The synthetic files contain more extremes in summer than in winter. The cause of this is not obvious since the entire procedure is symmetric about zero, i.e., minima and maxima should be roughly equivalent. Bootstrapping by itself cannot produce extreme values, since it is merely ‘re-shuffling’ what already exists. While the values in the TMY source files are recorded values, they are very unlikely to include an extreme value. This means that we are looking at extremes created by a convergence of bootstrapping and simulation of the SARMA model.

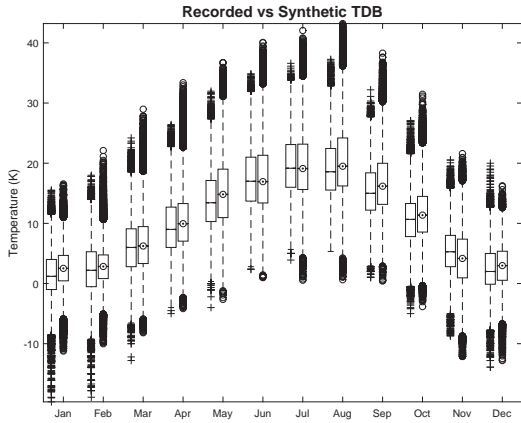


Figure 7: Recorded data are represented by box plots with lines for medians and crosses for outliers, synthetic data by targets and circles.

The difficulty here lies in assessing what an ‘acceptable’ extreme is: if 40°C has not been recorded in Geneva in the past 60-odd years, does that mean it will never happen in the future? Weather records of recent years say otherwise. As yet, we do not take a position on the matter and let the user decide on what they consider to be possible (though, highly improbable), or ridiculous. In the post-processing we removed points that were judged to be outliers using an arbitrary measure of their distance from the mean. A point of concern, not visible from these graphs, is that the synthetic data produces no ‘sequences of extremes’, i.e., heat or cold waves. This is being addressed in ongoing work.

Frequency distributions

Figures 6 and 7 suggest that the probability distributions of the synthetic series should have truncated left tails compared to the original data. That is to say they will not represent the winter extremes well enough. In fact, Figures 8 and 9 do show exactly that. Fig. 9 shows that the resampling procedure tends to smooth the unexpected peaks that exist in the Probability Distribution Function (PDF) of the TMY data (e.g., between about 5°C and 15°C). Unexpected in the sense that said peak is not present in the PDF of the recorded data. On the whole, the approximation is acceptably close to the original measured data.

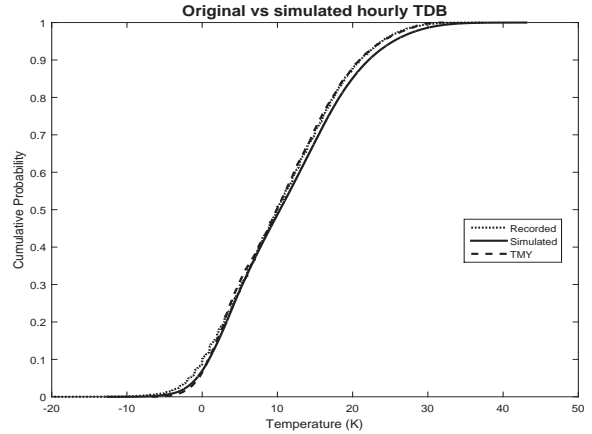


Figure 8: (Empirical) Cumulative Distribution Function (CDF) for synthetic, recorded, and TMY data from Geneva. The dotted line is for measured data, solid for synthetic, and dashed for TMY.

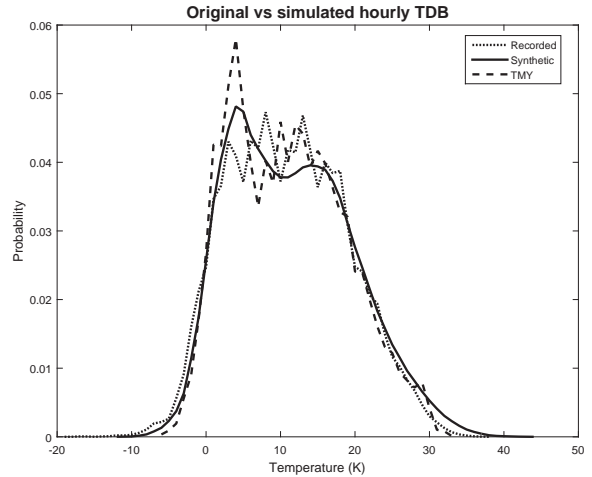


Figure 9: PDF of measured, synthetic, and recorded TDB data from Geneva. The dotted line represents measured data, solid represents synthetic, and dashed is for TMY.

Cross-Correlation

We proceed to examine the cross-correlations of temperature with solar radiation and humidity using Pearson’s r and Spearman’s ρ correlation coefficients. The quantity r measures the *linear* correlation between two variables, and ρ assesses how well the relationship between two variables can be represented as a monotonic function (Dodge 2008).

Table 2: Correlation coefficients for dry bulb temperature (TDB) with W and GHI. Pearson’s correlation coefficient is r , and Spearman’s is ρ .

	W		GHI	
	r	ρ	r	ρ
Original	0.83	0.86	0.51	0.47
TMY	0.84	0.86	0.52	0.44
Synthetic	0.79	0.83	0.41	0.4

The temperature predictably shows a strong linear correlation with humidity, which makes the high value of

Spearman’s coefficient unsurprising. Given that our original data was Relative Humidity (which depends on temperature), and the conversion to humidity ratio is an approximation based on the temperature and pressure, the two series are not necessarily independent. Strong correlation is also evident between Temperature and Solar Radiation. However, this might be a function of the relative sunniness of the example climate. Pending verifications for other, cloudier, climates, this particular relationship must be treated with caution. The correlations with both W and GHI are not appreciably different for the synthetic time series.

DISCUSSION

A crucial improvement from our previous efforts is that the ACF and PACF of the synthetic time series greatly resemble those of the original source material, the TMY, and also of measured data. Boland (1995) and Magnano et al. (2008) ascribe this to the fact that the mixing of “persistence effects at different time scales” (e.g., daily and seasonal moving averages or auto-regressive structures) is better represented by a model composed of separate parts explicitly representing these diverse time-scales. That is, separate models for annual and daily variability. The modelling procedure needs to be tested further, especially for climates that do not show a strong annual seasonality. For example, we found the residuals of the first-order Fourier fit to daily means for Mumbai to be quite different from those of Chicago and Geneva. The best outcome would be if the *structure* of the models is consistent across climates, and only the coefficients of the various polynomials changes. The next best is if the *structure* is consistent at least within a climate zone. The least generalisable outcome, of course, would be if the structure and coefficients are unique to each climate. If they turn out to be extremely sensitive to the initial data (i.e., the design reference file, or the ‘source’ file) then generating synthetic data will not alleviate the problem we set out to address: that each file is a ‘random’ sample and relying on any one is far too limiting. In this worst case, the synthetic data generated from each weather file will not be sufficiently representative of a reasonable range of climatic conditions to be considered a worthwhile alternative to simulation with a single file.

The choice of bootstrapping for re-creating synthetic series was motivated by the general applicability of the method and its relative insensitivity to underlying distributions. A notable set of cases for which the bootstrap is expected to fail is if the data come from a distribution which is in “the domain of attraction of a (non-normal) stable law” (Politis 1998). We work with bootstrapping raw material that has a t-distribution. While this doesn’t seem to be causing problems currently, this is a matter to be investigated further.

The most important limitation of any method based on fitting conditional mean and variance models to ob-

served data is that they are *temporary constructs*. We do not claim that our models are better representations of climate than global- and regional-level simulations of physical phenomena. It is incumbent upon the energy modeller who would like to use this strategy to check and recheck the model fits to ensure that crucial assumptions are not violated, and the synthetic time series are physically valid.

CONCLUSION

We have demonstrated the use of time series analysis to create synthetic weather data, particularly Fourier fitting, conditional mean models, and resampling. This synthetic weather ‘data’, while based on a constructed typical time series (TMY), turns out to be more representative of the full range of values seen in our example climate over the last sixty or so years. Ongoing work by the authors extends this approach to solar radiation and humidity. Ultimately, the use of synthetic weather data is a ‘brute force’ approach to the characterisation of uncertainty due to weather in buildings. That is, the only way to improve one’s estimate of bias in some statistical measure of model output is to create more paths (n_{sim} and n_{boot}) and simulate each resulting time series separately. We expect that the time required to improve one’s coverage would scale, at best, at an order of $\mathcal{O}(n)$. That is to say that 200 simulations should take twice the time as 100, up to some limits where post-processing the data becomes a problem. An approach based on stochastic differential equations seems far more elegant and was proposed more than 20 years ago by, for example, Haghghat et al. (1985). However, the available literature on this requires the implementation of stochastic differential equation solvers in building simulation software. Another approach of interest is the implementation of a stochastic weather generator *in* a building simulation programme, proposed by Scartezzini et al. (1990), which would obviate the necessity to store several weather files for simulation. However, this will not necessarily reduce the simulation time since one would still need a statistically representative sample of runs.

ACKNOWLEDGEMENT

The authors would like to thank Professor Anthony Davison, Mikael Kuusela, and Ben Miller. Their patient explanations, useful insights, and careful review have been invaluable in making the work possible. This work was supported in part by the EuroTech Universities Alliance and the CCEM SECURE Project.

Glossary

- n_{boot} Number of bootstrap samples.
- n_{sim} Number of simulations of the SARMA model.
- ACF** Auto-Correlation Function
- AIC** Akaike Information Criteria
- AR** Auto-Regressive

ASHRAE American Society of Heating, Refrigerating and Air-Conditioning Engineers
BIC Bayesian Information Criteria
CDF (Empirical) Cumulative Distribution Function
DRY Design Reference Year
GHI Global Horizontal Irradiation
leakage “The appearance of a non-zero value in the transform at a frequency f because of the presence of a sinusoid at a different frequency f_0 is called leakage.” (Bloomfield 2000)
MA Moving Average
MAD Median Absolute Deviation
PACF Partial Auto-Correlation Function
PDF Probability Distribution Function
PSD Power Spectral Density
RH Relative Humidity
SAR Seasonal Auto-Regressive
SARMA Seasonal Auto-Regressive Moving Average model
SMA Seasonal Moving Average
TDB Dry Bulb Temperature
TMY Typical Meteorological Year
W Humidity Ratio

REFERENCES

- Bhandari, Mahabir, Som Shrestha and Joshua New (2012). ‘Evaluation of weather datasets for building energy simulation’. In: *Energy and Buildings* 49, pp. 109–118.
- Bloomfield, Peter (2000). *Fourier analysis of time series: an introduction*. 2nd ed. Wiley series in probability and statistics. Applied probability and statistics section. New York: Wiley. 261 pp.
- Boland, John (1984). ‘Time Series Modelling of Solar Radiation’. In:
— (1995). ‘Time-series analysis of climatic variables’. In: *Solar Energy* 55.5, pp. 377–388.
- Boland, John, Jing Huang and Barbara Ridley (2013). ‘Decomposing global solar radiation into its direct and diffuse components’. In: *Renewable and Sustainable Energy Reviews* 28, pp. 749–756.
- Chinazzo, G. (2014). ‘Refurbishment of Existing Envelopes in Residential Buildings: assessing robust solutions for future climate change’. Master Thesis. Lausanne, Switzerland: EPFL.
- Chinazzo, G., Parag Rastogi and M. Andersen (2015). ‘Assessing robustness regarding weather uncertainties for energy-efficiency-driven building refurbishments’. In: 6th International Building Physics Conference. Torino.
- Crawley, DB and YJ Huang (1997). ‘Does it matter which weather data you use in energy simulations’. In: *User News* 18.1, pp. 25–31.
- Davison, A. C. (2003). *Statistical Models*. Cambridge University Press. 1331 pp.
- Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and their Application*. 1st. Cambridge University Press. 594 pp.
- Dodge, Yadolah (2008). *The concise encyclopedia of statistics*. 1st. ed. Springer reference. New York: Springer. 616 pp.
- Haghighat, F., T. E. Unny and M. Chandrashekar (1985). ‘Stochastic Modeling of Transient Heat Flow Through Walls’. In: *Journal of Solar Energy Engineering* 107.3, pp. 202–207.
- Hansen, James E. and Dennis M. Driscoll (1977). ‘A Mathematical Model for the Generation of Hourly Temperatures’. In: *Journal of Applied Meteorology* 16.9, pp. 935–948.
- Hong, Tianzhen, Wen-Kuei Chang and Hung-Wen Lin (2013). ‘A sensitivity study of building performance using 30-year actual weather data’. In: *Proceedings of BS2013*. 13th Conference of International Building Performance Simulation Association. Chambéry, France.
- IPCC Core Writing Team (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.
- Lund, Hans (1995). *The Design Reference Year Users Manual - A Report of Task 9 : Solar Radiation and Pyranometer Studies*. Copenhagen, Denmark.
- Magnano, L., John Boland and R. J. Hyndman (2008). ‘Generation of synthetic sequences of half-hourly temperature’. In: *Environmetrics* 19.8, pp. 818–835.
- MeteoSwiss (2014). *IDAWEB*. Federal Office of Meteorology and Climatology. URL: <https://gate.meteoswiss.ch/idaweb/system/welcome.do>.
- NCDC/NOAA (2014). *NNDC Climatic Data OnLine*. NOAA Satellite and Information Service: National Climatic Data Center (NCDC). URL: <http://www7.ncdc.noaa.gov/CDO/cdo>.
- NIST (2013). *NIST/SEMATECH e-Handbook of Statistical Methods*.
- Olsina, Fernando (2013). ‘Simulation of wind velocity for stochastic power market and power system analysis’. In: *Ingengerare revista de la facultad de ingeniería Pontificia Universidad Católica de Valparaíso* 27, p. 67.
- Politis, D.N. (1998). ‘Computer-intensive methods in statistical analysis’. In: *IEEE Signal Processing Magazine* 15.1, pp. 39–55.
- Rastogi, Parag and Marilyne Andersen (2013). ‘Generation of Weather Files Using Resampling Techniques: An Exploratory Study’. In: *Proceedings of BS 2013*. 13th International Conference of the International Building Performance Simulation Association. Chambéry, France, pp. 1422–1429.
- Scartezzini, J.-L., M. Nygard Ferguson and F. Bochud (1990). *Compression of multi-year meteorological data*. EF-REN (90) 009. Lausanne, Switzerland: Ecole polytechnique fédérale de Lausanne.