Renewable Energy 118 (2018) 343–350

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Development of robust meteorological year weather data

Sleiman Farah [a,*], Wasim Saman [a], John Boland [b]

[a] University of South Australia, School of Engineering, Barbara Hardy Institute, Adelaide, Australia
[b] University of South Australia, School of Information Technology & Mathematical Sciences, Barbara Hardy Institute, Adelaide, Australia

ARTICLE INFO

Article history:
Received 8 June 2017
Received in revised form
20 September 2017
Accepted 12 November 2017
Available online 14 November 2017

Keywords:
Typical meteorological year (TM2)
Robust synthetic data
Fourier series
Energy simulation
Levene's test
Kolmogorov-smirnov two-sample test

ABSTRACT

Building energy performance simulations are limited to typical meteorological weather conditions available in simulation software. Such simulations are insufficient for analysing energy performance sensitivity to a range of probable weather conditions. This research presents a method for developing robust meteorological weather data that can be used for energy performance sensitivity analysis without the need to access historical weather data. The method decomposes dry bulb temperature (DBT) and global horizontal solar radiation (H) into deterministic and stochastic components. For the typical weather data of the City of Adelaide, the deterministic component for each of DBT and H consists of a single frequency Fourier series. The stochastic components consist of 1-lag and 2-lags autoregressive models for DBT and H respectively. The stochastic components also include randomly selected values from the residuals of the autoregressive models. Based on this method, the coldest and hottest weather conditions were selected to simulate the energy performance of a single space. The results revealed 39% more cooling and 15% less heating in the hottest year, and 14% more heating and 64% less cooling in the coldest year. The results indicate that simulations based on typical weather conditions only are insufficient for assessing buildings' energy performance.

1. Introduction

Energy usage in the building sector accounts for approximately 20% of the total end-use of energy worldwide [1]. Reducing this high energy usage requires proper designs to improve the performance of buildings and buildings' energy systems. Typically, the designs are based on building energy simulation results to identify necessary size and capacity of different components. For instance, the results of such simulations can reveal whether additional thermal insulation for walls or roof is required. Such simulations consider ambient conditions and weather variables, such as dry bulb temperature (DBT), relative humidity and global solar radiation (H). Weather data, such as TM2 weather data, available in simulation software are compiled from historical and estimated weather data based on a defined statistical method [2]. Typically, a simulation software provides a single weather data file for a specific location which allows analysing a building's energy performance under typical long term weather conditions. Although designs based on typical long term weather conditions may seem suitable to achieve typical expected long term performance, such designs do not guarantee the expected performance when weather conditions deviate from that in the simulation. The uncertainties in weather conditions are often overcome by applying safety factors to oversize different components of the system. However, without proper analysis, this approach may unnecessarily oversize the system and increase the capital cost or undersize the system and reduce thermal comfort. In either of these two cases, the operating conditions of the system will not match those of optimal performance, which unnecessarily increases energy usage. Therefore, a proper analysis requires simulating the building's energy performance under multiple weather conditions which are as likely to happen as the typical conditions.

In addition, performance analysis of renewable energy and energy storage systems, as in Refs. [3,4], can be improved by using multiple weather conditions. Similarly, when the performance of a renewable energy system is optimised for a period more than a year [5,6], instead of repeatedly using the same weather conditions for the required number of years, a variation of likely weather conditions can be applied for different years. The different yearly weather conditions are likely to influence the optimisation results.

To analyse the performance sensitivity in relation to weather conditions, researchers can simulate the energy performance using historical weather data. However, this data is often unavailable and incomplete [7], and detailed processing is required to organise the

* Corresponding author.
E-mail address: sleiman.farah@mymail.unisa.edu.au (S. Farah).

https://doi.org/10.1016/j.renene.2017.11.033

data in a format, such as TM2, TM3 or EPW format, that is compatible with simulation software [8]. These limitations deny most researchers the ability to analyse building's performance sensitivity to weather conditions.

To overcome these limitations researchers can use weather data generators such as RUNEOLE [9]. Weather data generators are useful when weather data are unavailable. Using historical data, and physical and statistical models, weather data generators can generate typical and extreme weather conditions [10]. However, weather data generators are unnecessary when typical weather data are available, especially when the weather data generators may be unavailable to researchers.

When typical weather data are available, researchers can develop synthetic data that have similar statistical characteristics as the typical weather data. Synthetic weather data were generated in a top-down approach using CLIMED software. Using algorithmic chains, the software generated monthly weather data which were used as input to generate daily weather data, and the daily weather data were used to produce hourly data [7]. However, this process required accessing weather data similar to the considered site to adjust the model parameters used in the software. The generation of synthetic data were also explored by using the "smooth" function in MATLAB software to identify a trend in the weather data for each month, and the residuals between the identified trend and the initial data were then randomly resampled. The synthetic data from this process would be the sum of the monthly trends and the randomly resampled data. However, the generated synthetic data using this process were unsatisfactory as excessive fluctuations were observed in the trend for the entire year [11]. Based on a previous work of Boland [12], an improvement in identifying the trend was the use of Fourier series analysis [13]. Additional improvements were also included in the work of Rastogi and Andersen [13] for developing synthetic data based on typical weather data. These improvements included fitting a seasonal autoregressive moving average (SARMA) model to the Fourier series residuals and performing 3-days blocks of random sampling within each month to maintain the intrinsic weather inertia in the data as suggested by Magnano, Boland and Hyndman [14].

While these improvements produce synthetic data statistically similar to the original data, some of the adopted procedures complicate the process of synthetic data generation. For instance, instead of fitting a SARMA model to the Fourier series residuals, an autoregressive moving average (ARMA) model should be sufficient as all the important frequencies can be detected and detrended using the Fourier series analysis. In addition, using 3-days blocks for random sampling seems unnecessary as an ARMA model of the Fourier series residuals is meant to model the intrinsic inertia of weather data.

This research presents a method for developing a robust meteorological year (RMY), without the need of accessing historical weather data, based only on typical data available in simulation software. The method decomposes the data into deterministic (Fourier series) and stochastic (ARMA + residuals) components. The deterministic component is maintained the same throughout the process of developing the RMY data, while the stochastic component is modified by random sampling. This method has three main differences compared with other methods. First, this method is based on average daily values for the modelling of both deterministic and stochastic components. The average values smooth the data and simplify the modelling of deterministic and stochastic components. Second, instead of resampling each month separately, the method allows mixing the errors from different months. The mixing allows creating a wider range of variations in the robust data. Third, the generation of hourly data from average daily synthetic data is based on selecting hourly data from the initial data. This selection eliminates the need for detecting outliers in the hourly synthetic data.

The method presented in this research uses the TM2 weather data for the City of Adelaide developed by Meteonorm. The focus of the method is on DBT and H as the two main weather conditions affecting building energy performance.

2. Data deterministic component – Fourier series models

2.1. Dry bulb temperature

The seasonality of the daily average dry bulb temperature (DBT) is clearly shown in Fig. 1, with the average DBT in summer being higher than that in winter. The variation of DBT can be modelled by a deterministic function; a Fourier series (FS) which has a frequency equal to one cycle per year.

Higher frequencies that may represent quarterly (frequency equals 4) or monthly (frequency equals 12) variations could also be significant to represent the data with a FS. However, the frequency power spectrum shown in Fig. 2, reveals that the power corresponding to frequencies higher than the fundamental frequency (frequency equals 1) is negligible compared to the power of the fundamental frequency. This result indicates that using the fundamental frequency is sufficient to capture most of the periodic variation in the data.

Consequently, the FS model (T_{FS}) of the average daily DBT can be represented by Equation (1)

$$T_{FS} = \overline{T} + a\cos(\omega t) + b\sin(\omega t) \tag{1}$$

where $\overline{T}$ is the average temperature calculated as in Equation (2)

$$\overline{T} = \frac{\sum_{i=1}^{365} T_i}{365} \tag{2}$$

where T_i is the average daily DBT.

The value of ω is calculated as shown in Equation (3)

$$\omega = \frac{2\pi}{365} \tag{3}$$

and the remaining unknown coefficients a and b in Equation (1) are calculated to minimise the sum of the squared errors (SSE) between the data and the FS. The SSE is calculated as shown in Equation (4)

$$SSE = \sum_{i=1}^{365} \left(T_{FS_i} - T_i\right)^2 \tag{4}$$

The minimum value of SSE is achieved for a and b equal to 5.122 and 2.075 respectively, and the FS model can by written as in

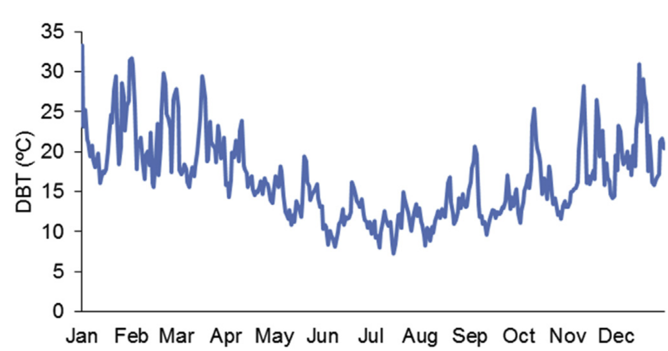

Fig. 1. Average dry bulb temperature (°C).