PLEA 2017 EDINBURGH

Design to Thrive

A new empirical model incorporating spatial interpolation of meteorological data for the prediction of overheating risks in UK dwellings

Matej Gustin[1, 2], Argyris Oraiopoulos[1, 2], Robert S. McLeod[1, 2] and Kevin J. Lomas[1, 2]

[1] Building Energy Research Group (BERG), School of Civil and Building Engineering, Loughborough University, Loughborough, UK;

[2] London – Loughborough Centre for Doctoral Training in Energy Demand (LoLo CDT), Loughborough University, Loughborough, UK, a.oraiopoulos@lboro.ac.uk

Abstract: Heat-related morbidity and mortality is anticipated to increase as climatic change induced overheating become increasingly common. The development of building-specific predictive models has the potential to alert occupants and emergency services to the severity of impending risks. This research aims to evaluate the implementation of a newly developed time series model for overheating prediction. Since risk forecasting is contingent upon the accuracy of the model at different future time steps, the sensitivity of model outputs to the uncertainty in the data inputs needs to be understood. Internal and external climatic variables were monitored in an unoccupied domestic dwelling in order to evaluate the empirical model's predictive accuracy. The uncertainty related to the proximity of external weather stations was evaluated using data taken from four nearby weather stations and further bespoke data sets derived by interpolation. The results confirmed the overall accuracy of the newly developed time series predictive model, whilst highlighting the benefits of climatic data interpolation in reducing predictive uncertainties. The empirically derived modelling approach showed a low variance to the actual temperature evolution over a seven-day predictive period, pointing to its validity as a robust model for the prediction of future overheating risks.

Keywords: Overheating, Time Series Analysis Method, Uncertainty, Meteorological data, Spatial interpolation

Introduction

Context

Time series data, spanning more than a century, illustrates that the most populous regions on Earth are experiencing progressively hotter annual temperatures (NASA & GISS, 2017). Warmer than average summers coupled with an increased frequency of extreme heat wave events (Jenkins et al., 2008) pose obvious risk factors in relation to overheating in the built environment. Events such as the 2003 heat wave, which is reported to have resulted in over 2000 heat-related deaths in the UK, are predicted to become increasingly common (Armstrong et al., 2011; Hajat et al., 2006; Rooney et al., 1998; Wright et al., 2005).

Active cooling systems remain relatively uncommon in UK homes but their uptake is projected to increase rapidly (Pathan et al., 2008). O'Neill et al. (2005) point out that unless extensively subsidised however, the ownership of cooling systems is likely to reflect socioeconomic inequalities thereby rendering disadvantaged households more vulnerable. Furthermore widespread power outages in North America and Australia during heat waves have highlighted the risks associated with reliance upon active cooling systems at times of grid overload (Ostro et al., 2010).

Despite strong epidemiological correlations between elevated external temperatures and increased risks of heat-related morbidity and mortality (Armstrong et al., 2011; Vardoulakis & Heaviside, 2012) relatively little is known about internal temperature evolution under these conditions. Real-time predictive overheating models are needed in order to understand when critical thresholds are likely to be breached in specific buildings and to warn occupants and facility managers when health-endangering environmental conditions are anticipated to occur.

Background – the development of time series models and indoor temperature prediction

Studies related to overheating in dwellings, conducted in recent years, can be broadly categorised as: those that involved measuring internal temperatures in order to identify and quantify the risk of overheating (Beizaee et al., 2013; McLeod and Swainson, 2017; Pathan et al., 2017), those that involved dynamic thermal simulation modelling to assess the current and future risk of overheating in dwellings (McLeod et at., 2013; Mavrogianni et al., 2012); and those that have used empirical data to construct predictive models in order to assess overheating both spatially and temporally (Mlakar & Strancar, 2011; Mirzaei et al., 2012). Modelling methods that make use of measurements to explain the variation in the data present advantages over other modelling methods that typically entail large numbers of assumptions and thereby elevate the level of uncertainty in the results. The Time Series Analysis Method (TSAM) has been successfully used in diverse fields such as economics, geophysics, control engineering and meteorology to describe, explain, predict and control processes (Chatfield, 1996). Here the descriptive TSAM is applied to room temperature data. The advantage of TSAM over other black box approaches (such as artificial neural networks, autoregressive models, multiple linear regression models, distributed lag models, transfer functions etc.) is that it is an exploratory technique that allows a clear understanding of the causality of internal temperatures.

Methodology

An empirical TSAM model was used to predict the hourly internal temperature evolution in the test house, in relation to different sources of meteorological data. The data measured on-site were compared with the meteorological data and the modelling predictions in order to assess the effects of the proximity of weather stations and interpolation of weather data on the models' predictive accuracy.

Time series model

This paper utilises a newly developed Internal Trend and Cyclical Component (ITCC) Model, which is based on the descriptive TSAM approach. The principle aim of the descriptive TSAM is to decompose the variation in the series into individual components (Trend, Cyclical variation) that can be described and modelled independently (Fleming & Nellis, 1994). The definition of these components is specific to the dataset used in the analysis. In this work, the components *trend* and *cyclical variation* are defined in relation to the internal air temperature profiles in homes. The *trend* represents the changes in the series from day to day over the (52 days) monitoring period which encompasses the daily variation. Whilst the *cyclical variation* encompasses the diurnal fluctuations of internal temperatures (the variation within a 24h period). These concepts have their origins in the formative work of Chatfield (Chatfield, 1996) and Kendal & Ord (1990).

The detailed methodology describing the ITCC model can be found in Oraiopoulos et al. (2017). The ITCC model is the result of joining the *trend* and the *cyclical* component models together and is given below:

$$\boldsymbol{\theta'_{in,t} = i + g \times \left[(1-\alpha) \times \left(\overline{\theta}_{ex,d} + \alpha\,\overline{\theta}_{ex,d-1} + .. + \alpha^m\,\overline{\theta}_{ex,d-m}\right)\right] + A \times C_{ex,\varphi_e} + B \times C_{s,\varphi_s} - \gamma}$$

Where:

$\theta'_{in,t}$	hourly modelled internal air temperature
i	y-axis intercept of the line of best fit for the correlation between the daily mean internal temperature and the exponentially weighted moving average of the daily means external air temperatures
g	gradient of the line of best fit for the correlation between the daily mean internal air temperature and the exponentially weighted moving average of the daily means of the external air temperatures
α	constant between 0.00 and 1.00 (BS EN 15251, 2007)
$\overline{\theta}_{ex,d}$	daily mean of the external air temperature of the current day
$\overline{\theta}_{ex,d-1}$	daily mean of the external air temperature of the previous day
m	total number of previous days used in the formula for the exponentially weighted moving average of the daily mean of the external air temperature
$\overline{\theta}_{ex,d-m}$	daily mean of the external air temperature of the m^{th} previous day
A	numerical coefficient of the cyclical component of the external air temperature
C_{ex,φ_e}	cyclical component of the external air temperature
φ_e	phase of the cyclical component of the external air temperature
B	numerical coefficient of the cyclical component of the solar irradiation
C_{s,φ_s}	cyclical component of the solar irradiation
φ_s	phase of the cyclical component of the solar irradiation
γ	Constant

Test house and empirical data monitoring

The data used in this study was collected from one of two unoccupied, semi-detached test houses located in Mountsorrel, Leicestershire. The experiment was undertaken, and all the data collected, by other researchers as part of a different research project (see acknowledgements), during the summer of 2016. The houses are typical, solid brick walled, family homes dating from the 1910s which have subsequently undergone some refurbishment.

Figure 1. Front elevation (facing S-SE) of the test houses (the left hand house was used in this study).

Weather and environmental data

Internal monitored data

Internal temperature data was collected in the test house from 9 June to 31 July (2016) using a U-type thermistor (accuracy = ±0.2°C) installed in the middle of each room (with an aluminium foil shielding to mitigate radiant heat influences). Logging was carried out at five-minute intervals. During the tests, the windows remained closed, with blinds and trickle vents open. Synthetic occupancy was created, simulating the occupancy profile and appliance usage of an elderly couple (staying at home the whole day): living room (occupied from 08:30 to 23:00) and bedroom (occupied from 23:00 to 07:30).

External weather data – measured on site

External temperatures were monitored on site during the monitoring period (from 9 June to 31 July) using a shielded U-type thermistor (accuracy = ±0.2°C). The sensor was mounted in a shaded location on the North side of the house. The data was logged at the same interval as the internal sensors (five-minute intervals).

External weather data – third party sources

Hourly weather data was gathered from two different sources: Loughborough University (LU) meteorological station and the Centre for Environmental Data Analysis (CEDA, n.d.). Through the CEDA platform, it was possible to access data from the Met Office Integrated Data Archive System (MIDAS): UK Hourly Weather Observation Data. Three MIDAS stations in the proximity of the test house were selected for this study: Sutton Bonington (SB), Coundon Coventry (CC) and Wittering (WIT). The locations of the test house, meteorological stations and their distances from the site are shown on the map in Figure 2.

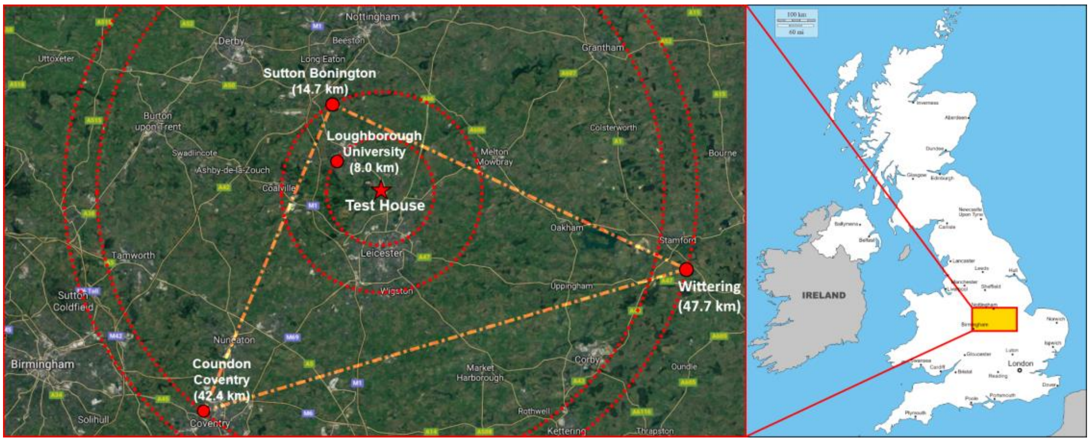

Figure 2. Locations of the test house (star) and meteorological stations (circles) on the map.

External weather data – spatial interpolation

For the spatial interpolation of meteorological data, the method adopted by the Joint Research Centre (JRC) of the European Commission was selected (Voet, Diepen, & Voshaar, 1994). This method was chosen due to its proven reliability and ease of application. Voet et al. (1994) demonstrated that with the averaging of data from optimally sited meteorological stations it is possible to obtain satisfactory results without the use of weighting coefficients or other more complex interpolation methods such as splines, kriging etc. (Franchello, 2005; Hofstra et al., 2008). Although the JRC method was originally developed for interpolating daily meteorological data, in this study the same criteria were adopted for spatial interpolation at an hourly resolution. As suggested by Voet et al. (1994), the optimal number of three weather stations for interpolation was used in this study. The results of interpolation using two meteorological stations are also shown for reference. In this case study, no algorithm was adopted for their selection, but the three closest meteorological stations (triangulating the site) with complete hourly temperature and solar radiation data were chosen. The JRC method foresees that in order to improve the accuracy of the data, the air temperature and has to be corrected to take into account the altitude difference. For every 100m increase in altitude (relative to the site) the temperature was reduced by 0.6°C. To interpolate the data, the corrected station data was then simply averaged.

Overheating

Various criteria have been developed to assess when rooms in a dwelling might be considered as overheated. These include CIBSE static criteria which suggest that the operative temperature (OT) in living rooms should not exceed 28^{o}C for more than 1% of occupied hours in the year, for bedrooms the criterion is 1% of hours over 26^{o}C (CIBSE, 2006). More recently a move towards the use of adaptive overheating thresholds, which vary according to the outdoor temperature, have gained popularity for the assessment of risks in free running (i.e. naturally ventilated) buildings (CIBSE, 2013). In terms of legislation The Health and Safety Rating System (HHSRS) Operating Guidance states that when temperatures exceed 25°C there is a significant increase in the risk of strokes and mortality (HHSRS, 2004). Currently the HHSRS provides the only statutory definition of 'overheating' risks in relation to morbidity and mortality in UK residential properties (McLeod and Swainson, 2017).

Uncertainties analysis

As stated by Hopfe et al (2013), "in the assessment of the performance of a building, it is imprudent to take deterministic values for the input parameters". Moreover, to generate robust predictions, analysis of the measurement uncertainties is required (Buswell, 2013). In order to take measurement uncertainties into account, they were evaluated in accordance with good practice guidance developed by the National Physics Laboratory (Bell, 2001). Using type A (repeated readings) and type B (manufacturer specifications – accuracy) uncertainties for the air temperature measurements were calculated to be ±0.10°C and ±0.12°C respectively, producing a combined standard uncertainty of ±0.15°C. Considering a coverage factor of k=2, the resulting extended standard uncertainty is ±0.30°C, with a confidence interval of 95%.

Model evaluation

The ITCC time series model was trained from 9 June to 16 July, and produced hourly predictions from 17 to 31 July. Since the external air temperature was also monitored on-site, in order to assess the data from the various meteorological stations and the interpolated data, the RMSE between the measured and adopted meteorological data was used as a dispersion metric, as used for example by (Voet et al., 1994). The RMSE was also used to evaluate the errors between predicted and measured internal temperatures in the living room and bedroom. To check the influence that measurement uncertainty has on the results, the RMSE and R^2 were also calculated on the measurement uncertainty limits as an adjunct to using the deterministic values. These values represent the minimum errors and the maximum explanatory power of the model (respectively), which could be achieved when considering the uncertainty in the measurements. Since the model was predicting reliably for the period up to seven days, the errors between the predicted and modelled data were evaluated only for the first week of predictions.

Results and Analysis

Modelling inputs – weather data

As shown in figure 3, when the meteorological data is taken from a single station, the RMSE ranges between 0.96 and 1.15°C, compared to the site values. The RMSE is significant also for the meteorological stations that are very close to the test house. It has to be mentioned that the errors for the MIDAS stations are at certain hours very large, suggesting that the air temperatures are varying more quickly than at the analysed site. This can be easily explained due to the locations of the MIDAS stations, which are usually located in open field, whilst the test house is located in the middle of a small town. However, when the meteorological data is interpolated, the RMSE drops significantly. The best results were obtained using triangulation (interpolation of three meteorological stations), with the RMSE ranging between 0.69 and 0.89°C. This indicates that using triangulation it is possible to improve the input data by 7-28% relative to using a single 'near neighbour' weather station.

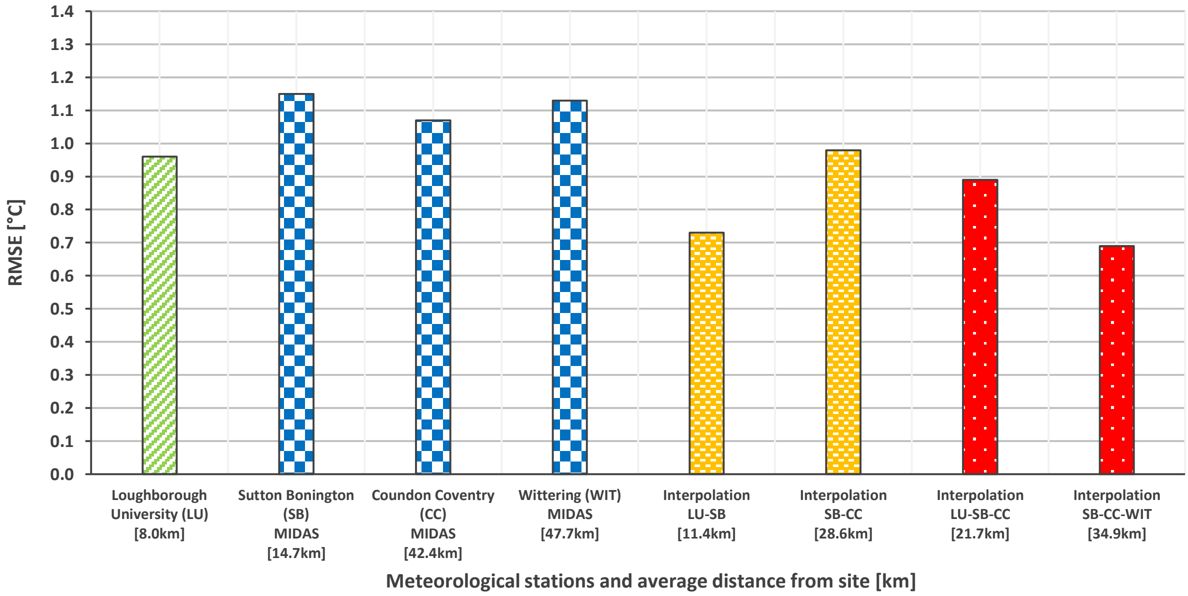

Figure 3. RMSE of the meteorological data against the external temperatures measured on site.

Modelling outputs – hourly predictions

Figure 4 illustrates how the model was trained on 38 days of data to predict the internal temperature evolution for the following 14 days. Due to the increasing inaccuracy of forecasts after day 7, only the first week of predictions was considered in this study.

Overall, the model showed an excellent explanatory power for the first week of predictions for all the meteorological data, with an R2 of 0.876-0.896 for the living room (LR), and 0.943-0.952 for the bedroom (BR). Considering the measurement uncertainty, the explanatory power might be even higher with the potential maximum of 0.937-0.951 (LR) and 0.951-0.960 (BR). As shown in figure 5, taking the data from the closest meteorological station does not guarantee that the most accurate prediction will be achieved. Indeed, in this case, the predicted data showed that while the largest RMSE was obtained from the two closest stations (LU and SB), the lowest RMSE (CC and WIT) was obtained from the two furthest stations. Whilst the interpolation did not show significant reductions in the RMSE of predictions, it is evident that with its use the errors show a greatly reduced variability (Figures 3 and 5), thereby reducing the uncertainty of the data that are used for the model.

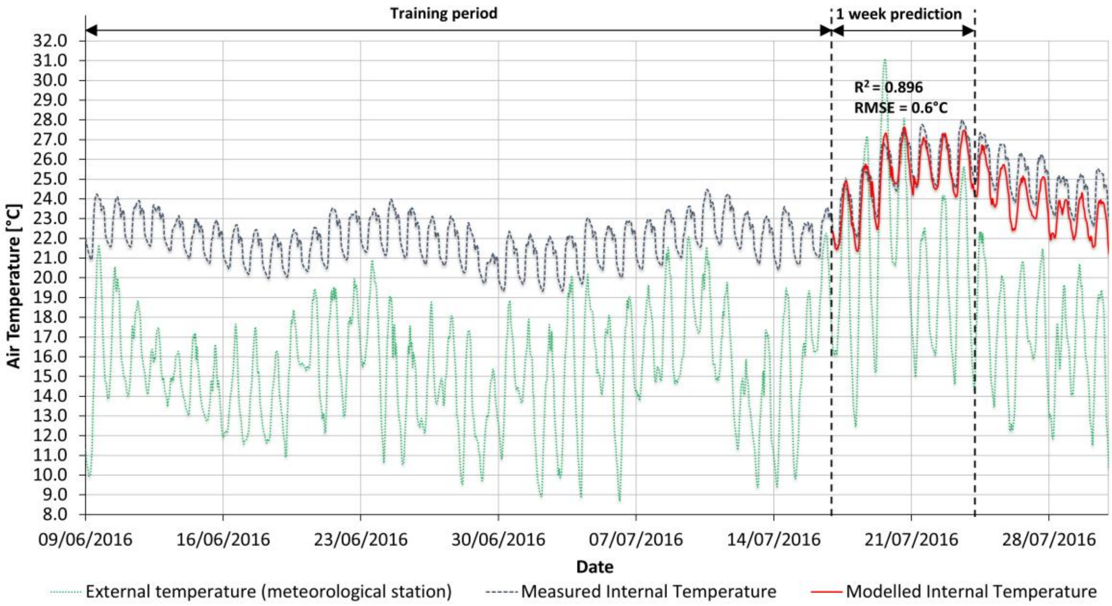

Figure 4. Measured and modelled temperatures – Example showing model dry bulb temperature predictions for the living room using meteorological interpolation (SB-CC-WIT)

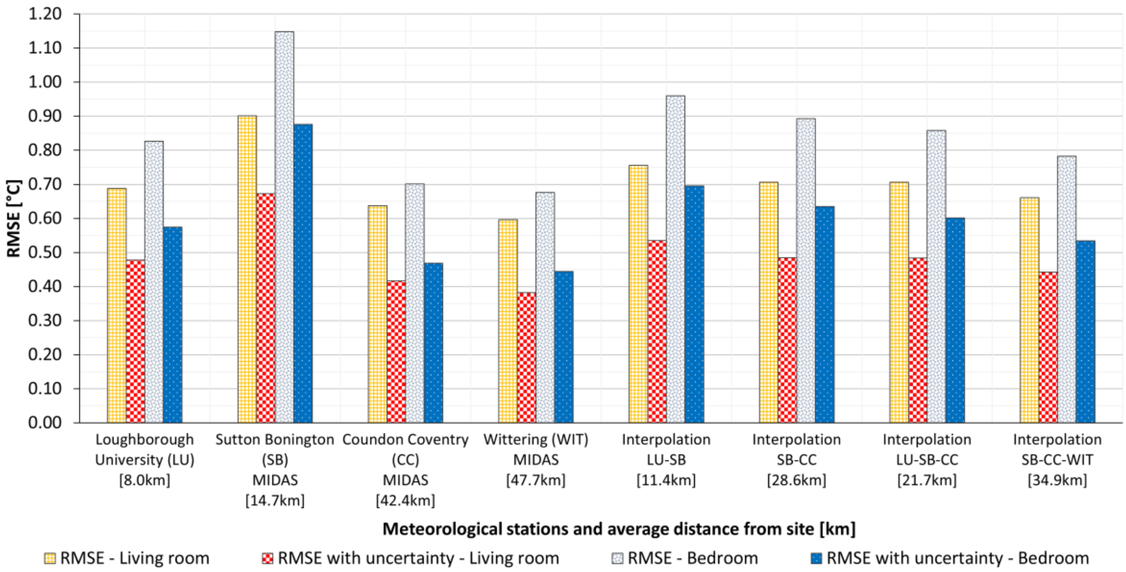

Figure 5. RMSE of the predictions in the LR and BR, with and without considering the uncertainty range.

Conclusions and recommendations

This study has shown that the newly developed ITCC empirical TSAM model is able to accurately predict the internal temperature evolution of a dwelling, for a period up to 7 days in the living room and bedroom of this test house. It also highlights the inaccuracies that are introduced to the model when data from 'near neighbour' meteorological stations are used. Triangulation of the weather data inputs improved the model's predictive accuracy whilst reducing the variability and uncertainties associated with the results. For more robust and prolonged predictions, further model development is needed in order to further improve predictive accuracy during sudden spikes in the external temperatures.

Acknowledgments

This research was made possible by EPSRC support for the London-Loughborough CDT in Energy Demand (grants EP/L01517X/1 and EP/H009612/1). The authors would like to thank

Dr. Stephen Porritt, Ms. Vicki Tink, Dr. David Allinson and Prof. Dennis Loveday for sharing the data used in this analysis.

References

Armstrong, B.G. et al. (2011). The association of mortality with high temperatures in a temperate climate: England and Wales. *Journal of Epidemiology & Community Health*, 65(4), pp.340–345.

Beizaee, A., Lomas, K.J. & Firth, S.K. (2013). National survey of summertime temperatures and overheating risk in English homes. *Building and Environment*, 65, pp.1–17.

Bell, S. (2001). Measurement Good Practice Guide No. 11- A Beginner's Guide to Uncertainty of Measurement.

BS EN 15251 (2007). Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality , thermal environment , lighting and acoustics.

Buswell, R. (2013). Uncertainty in Whole House Monitoring. In *13th IBPSA conference*, Chambery, France, 2013.

CEDA (Centre for Environmental Data Analysis). Retrieved from http://www.ceda.ac.uk/.

Chatfield, C. (1996). *The analysis of time series : An introduction*,5th ed., London: Chapman & Hall.

CIBSE (2006). *CIBSE Guide A: Environmental Design*.

CIBSE TM52 (2013). The limits of thermal comfort : avoiding overheating in European buildings.

Fleming, M.C. & Nellis, J.G. (1994). *Principles of applied statistics*, London: Routledge.

Franchello, G. (2005). Assessment of Interpolation Methods for EFAS, IES (JRC).

Hajat, S. et al. (2006). Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology*, 17(6), pp.632–638.

HHSRS (2004). Housing Health and Safety Rating System.

Hofstra, N. et al. (2008). Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research Atmospheres*, 113(21).

Hopfe, C. J., Augenbroe, G. L. M., & Hensen, J. L. M. (2013). Multi-criteria decision making under uncertainty in building performance assessment. *Building and Environment*, *69*, 81–90.

Jenkins, G.J., Perry, M.C. & Prior, M.J. (2008). The climate of the United Kingdom and recent trends.

Kendall, M. & Ord, J.K. (1990). *Time-series* 3rd ed., London: Edward Arnold.

Mavrogianni, A. et al. (2012). Building characteristics as determinants of propensity to high indoor summer temperatures in London dwellings. *Building and Environment*, 55, pp.117–130.

McLeod, R.S., Hopfe, C.J. & Kwan, A. (2013). An investigation into future performance and overheating risks in Passivhaus dwellings. *Building and Environment*, 70, pp.189–209.

Mcleod, R.S. & Swainson, M. (2017). Chronic overheating in low carbon urban developments in a temperate climate. *Renewable and Sustainable Energy Reviews*, 74 (September 2016), pp.201–220.

Mirzaei, P.A. et al. (2012). Indoor thermal condition in urban heat Island - Development of a predictive tool. *Building and Environment*, 57, pp.7–17.

Mlakar, J. & Štrancar, J. (2011). Overheating in residential passive house: Solution strategies revealed and confirmed through data analysis and simulations. *Energy and Buildings*, 43(6), pp.1443–1451.

NASA & GISS (2017). Global Temperature. *Climate Change: Vital Signs of the Planet*.

O'Neil, M.S., Zanobetti, A. & Schwartz, J., 2005. Disparities by race in heat-related mortality in four US cities: the role of air conditioning prevalence. *Journal of Urban Health*, 82(2), pp.191–7.

Oraiopoulos, A. et al. (2017). Development of a Statistical Model for the Prediction of Overheating in UK Homes Using Descriptive Time Series Analysis. In *15th IBPSA conference*, San Francisco, USA, 2017.

Ostro, B. et al., 2010. Effects of Temperature and Use of Air Conditioning on Hospitalizations.*American Journal of Epidemiology*, 172(9), pp.1053–1061.

Pathan, A. et al. (2017). Monitoring summer indoor overheating in the London housing stock. *Energy & Buildings*, 141, pp.361–378.

Pathan, A., Young, A. & Oreszczyn, T. (2008). UK domestic air conditioning: a study of occupant use and energy efficiency. In Windsor: Network forComfort and Energy Use in Buildings.

Rooney, C. et al. (1998). Excess mortality in England and Wales, and in Greater London, during the 1995 heatwave. *Journal of Epidemiology & Community Health*, 52, pp.482–6.

Vardoulakis, S. & Heaviside, C. (2012). *Health Effects of Climate Change in the UK 2012: Current evidence, recommendations and research gaps*, Health Protection Agency.

Voet, P. van der, Diepen, C.A. van & Voshaar, J.O. (1994). Spatial interpolation of daily meteorological data: A knowledge-based procedure for the region of the European Communities.

Wright, A., Natarajan, S. & Young, A. (2005). Dwelling temperatures and comfort during the August 2003 heatwave. *Building Services Engineering Research and Technology*, 26(4), pp.285–300.