**Forecasting indoor temperatures during heatwaves using time series models**

This item was submitted to Loughborough University's Institutional Repository by the/an author.

**Citation:** GUSTIN, M., MCLEOD, R.S. and LOMAS, K.J., 2018. Forecasting indoor temperatures during heatwaves using time series models. Building and Environment, 143, pp. 727-739.

**Additional Information:**

- This is an Open Access Article. It is published by Elsevier under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

**Metadata Record:** [https://dspace.lboro.ac.uk/2134/34595](https://dspace.lboro.ac.uk/2134/34595)

**Version:** Published

**Publisher:** © The Authors. Published by Elsevier Ltd.

**Rights:** This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Forecasting indoor temperatures during heatwaves using time series models

Matej Gustin\textsuperscript{a,b,*}, Robert S. McLeod\textsuperscript{a,b}, Kevin J. Lomas\textsuperscript{a,b}

\textsuperscript{a} School of Architecture, Building and Civil Engineering, Loughborough University, LE11 3TU, UK
\textsuperscript{b} London-Loughborough (LoLo) EPSRC Centre for Doctoral Training in Energy Demand, Loughborough University, LE11 3TU, UK

A R T I C L E   I N F O

Keywords:
Time series forecasting
Machine learning
Black-box model
ARX model
ARMAX model
Overheating

A B S T R A C T

Early prediction of impending high temperatures in buildings could play a vital role in reducing heat-related morbidity and mortality. A recursive, AutoRegressive time series model using eXogenous inputs (ARX) and a rolling forecasting origin has been developed to provide reliable short-term forecasts of the internal temperatures in dwellings during hot summer conditions, especially heatwaves. The model was tested using monitored data from three case study dwellings recorded during the 2015 heatwave. The predictor variables were selected by minimising the Akaike Information Criterion (AIC), in order to automatically identify a near-optimal model. The model proved capable of performing multi-step-ahead predictions during extreme heat events with an acceptable accuracy for periods up to 72 h, with hourly results achieving a Mean Absolute Error (MAE) below 0.7°C for every forecast. Comparison between ARX and AutoRegressive Moving Average models with eXogenous inputs (ARMAX) models showed that the ARX models provided consistently more reliable multi-step-ahead predictions. Prediction intervals, at the 95% probability level, were adopted to define a credible interval for the forecast temperatures at different prediction horizons. The results point to the potential for using time series forecasting as part of an overheating early-warning system in buildings housing vulnerable occupants or contents.

1. Introduction

1.1. Background

Overheating in homes and residential care facilities is increasingly acknowledged as a serious problem for developers, property owners/managers, landlords, tenants, health care providers and policy makers [1–3]. Climate change projections indicate that the majority of the world’s most populated regions will experience more frequent and more intense heat wave periods over the coming decades [4,5]. Warmer than average summers coupled with an increased frequency of extreme heat wave events [6] pose obvious risk factors in relation to overheating in the built environment.

By 2040 average summertime temperatures in the UK are expected to reflect those experienced during the heatwave of 2003 [7,8]; an event which caused over 2000 heat-related deaths in the UK and more than 30,000 across Europe [9]. Those most affected by excess heat are the elderly (over the age of 60 years), who are at an increased risk of heat-related illness [10] with those over the age of 65 years having a higher risk of heat-related mortality [11]. Because of the rising average life expectancy in the UK and other developed nations [12], premature mortality rates are anticipated to increase when similar events occur in the future.

It is well established that in a temperate climate high ambient temperatures are associated with an increased mortality rate [15,16]. In response to growing concerns regarding global warming and the predicted increase in the frequency of extreme heat-related events, heat plans have been created in a number of countries worldwide in order to establish a collaboration between meteorological services, civil protection and public health authorities to inform and protect residents from the impending health risks of hot weather [11]. Such plans are known as Heat-Health Warning Systems (HHWSs) and are activated when the weather is forecasted to exceed certain thresholds that might lead to adverse health effects. HHWSs are intended to make emergency responses more efficient and better coordinated in order to reduce heat-related morbidity and mortality [12]. It is well known that people in developed countries tend to spend most of their time indoors [13], and it has been established that even healthy people situated indoors are at a significantly higher risk of experiencing adverse conditions of extreme heat than the same healthy individuals would be if they were located outdoors [14]. Since HHWSs are based solely on the outdoor weather conditions and because these warnings are triggered on a regional or national level, it is impossible to accurately identify which dwellings or people are actually at risk.

Excess heat-related deaths can largely be attributed to respiratory...
diseases and cardiovascular causes, including: strokes, coronary heart diseases and congestive heart failures [13,14]. However, a study by Rooney et al. [15] observed that mortality during heatwaves occurring late in the summer is lower than at the beginning of the summer, a finding which suggests that seasonal acclimatisation processes may increase resilience to heat stress. People living in different regions, cities, urban and rural areas are accustomed to different temperatures and respond to heat differently [16]. Coupled with the fact that indoor thermal conditions do not depend solely on the external weather conditions, but also on the building characteristics and occupant behaviour, it is clear that associating heat-related risks exclusively with external temperature thresholds at a regional or national level is inadequate and that the development of local, dwelling-based thresholds, should be a priority [17].

Despite strong epidemiological correlations between elevated external temperatures and increased risks of heat-related morbidity and mortality [18], relatively little is known about the health impacts in residential buildings, and this has been identified as an area requiring further research [19]. Since the overheating criteria currently used in the built environment are based on thermal comfort and not health, the definition and incorporation of heat-safety metrics are required to assess heat stress in dwellings [19]. The positive correlation between core body temperatures and indoor temperatures [20], points to the potential of developing indoor health indices based directly on indoor temperatures. With the development of dwelling-based indoor thresholds that would associate heat-related risks directly to the indoor environmental conditions, predictive models could play an important role as part of a real-time warning device that would allow the timely detection of critical indoor thresholds. Knowing when such thresholds were likely to be breached in specific spaces would allow carers or facility managers to warn occupants when health-endangering environmental conditions are expected to occur. If widely deployed, such a system could help to avoid or reduce heat-related morbidity and mortality occurring during hot weather conditions and in the case of dwellings with vulnerable occupants, such a system could trigger the prompt dispatch of the emergency services.

Recent studies related to overheating in dwellings can be broadly divided into three categories: firstly, studies that have involved measuring internal air temperatures (and other physical variables) in order to identify and quantify the risk of overheating [21–24]; secondly, those that involved either quasi-steady state or dynamic thermal simulation modelling to assess the current and future risk of overheating [23,25–28]; and lastly, studies that have used empirical data to construct forecasting models for the prediction of the indoor thermal conditions [29–33]. The availability of observed data from large monitoring studies [21–23,28,34] provides the potential to develop empirical models which make predictions based on the data alone (i.e. machine learning). Machine learning models (which lack explicit physical parameterisation) are often referred to as black-box models or more generally as statistical models [35,36]. Such models have the potential to be used in forecasting the short-term future internal temperatures in buildings based solely on the external climate data and previously recorded internal temperatures. As such, black-box models are computationally and resource efficient and do not require any physical information describing the room or building fabric. If proven reliable, such models could be usefully deployed to protect building occupants from the impending risks of overheating in a specific space. Provision of tailored information to occupants (or their carers) and/or facilities managers advising on the level of preventative action needed to mitigate the risks is then possible.

1.2. Modelling methods

Different types of black-box models can be adopted for the prediction of internal air temperatures, with the most common being Time Series and Artificial Neural Network (ANN) models [37]. Whereas simpler time series forecasting models are based on linear methods, ANN allows more complex non-linear relationships between the response variable and its predictors [38]. Nevertheless, ANN models are harder to train, require large amounts of learning data and are limited by their lack of interpretability [36]. In addition, ANN models are known to give different results after repeated trials on the same data [30]. For these reasons, linear time series models offer several advantages over their non-linear ANN counterparts: namely that they are simpler to deploy, and the same data and inputs will always produce the same model parameterisation [30].

Ashiani et al. [32] observed that their time series regression model was not able to forecast accurately during a heatwave event, with the best model achieving an $\text{RMSE}$ of 2.10 °C, which points to the difficulty of developing a reliable forecasting model that is able to generalise with acceptable accuracy during such extreme events. Time series forecasting models such as AutoRegressive models with eXogenous inputs (ARX) and AutoRegressive Moving Average models with eXogenous inputs (ARMAX) have been shown to provide reasonably accurate short-term forecasts (2–3 h) of internal temperatures in office buildings when using high-resolution data (i.e. 5–15 min sampling); such models have achieved a Root Mean Square Error ($\text{RMSE}$) of 0.33 °C for 3-h forecasts using ARX models [33] and a Mean Absolute Error (MAE) of 0.11–0.19 °C for a 2-h forecasts using ARMAX models [30]. However, in these studies [30,33], forecasts were made for relatively mild summer days without sudden temperature spikes, and with peak indoor temperatures of approximately 24 °C [30] to 26 °C [33]. Although the forecasting accuracy of the models developed in the studies was good [30,33], they have been primarily developed to improve HVAC system control in air-conditioned offices and schools where there is detailed information available regarding their operation. As such their use cannot be directly transposed to free-running dwellings where HVAC systems are not generally used or to dwellings where space conditioning is used but operation schedules are far more unpredictable than in offices. Moreover, the use of sub-hourly data which HVAC system control development is predicated upon assumes that there is widespread availability of such high-resolution data and weather forecasts. In reality, if the models were to be integrated as a part of an indoor heat warning device, and deployed on a large scale, using forecasted weather data from national meteorological services, such as the UK Met Office, weather forecasts are not available at a sub-hourly resolution. Consequently, there is a lack of literature in relation to the development and prototyping of predictive models in relation to forecasting indoor temperatures over extended forecasting horizons that are able to operate at an hourly data resolution in free-running dwellings and with acceptable forecasting accuracy during extreme hot spells.

It is well known in forecasting that predictions are difficult to perform where the values of future predictors fall outside the range of the past (training) values [38]. Hence, if models are not validated over a suitably hot spell it is uncertain whether they would be able to forecast accurately during a heatwave.

In a study by Rios-Moreno et al. [29], it was found that for predictions of the internal temperature, ARX models generally performed more accurately than ARMAX models. Nonetheless, only one-step-ahead forecasts were performed in this study. On the other hand, Mustafaraj et al. [31] found that ARX and ARMAX models produced similar results, with ARMAX models being preferable for multi-step-ahead forecasts. Hence it is unclear from the literature which type of model is capable of providing the most accurate and consistent predictions of internal temperatures during heat waves, especially during extended forecasting horizons.

When longer forecasting horizons (h) are desired, multi-step-ahead forecasts (i.e. advanced predictions of multiple time steps) are required.
For this purpose, either recursive or direct strategies can be adopted [39]. In a direct strategy, a separate model is trained and adopted for each forecasting horizon [40], using similar methods to those used for one-step-ahead predictions, but with a longer prediction step [41]. In a recursive strategy, the prediction from a one-step-ahead model is used as an input for future prediction horizons [39] and subsequent predictions are performed by simply reiterating short-term predictors [41]. Whereas direct strategies might be superior for mis-specified models (i.e. when the considered class of models is sub-optimal and fails to account for the relationship between the explanatory and response variables), a recursive strategy may be better suited for well-specified models [40] and for long-range forecasting [42]. In addition, a huge advantage of the recursive strategy is that only one model is required, which considerably reduces the computational time, especially when many inputs are adopted and continuous long-range forecasting outputs are required [40]. Conversely, recursive forecasting is known to produce biased predictions when the underlying model is non-linear, particularly at longer horizons [40]. This is because any uncertainty or error generated in the one-step-ahead prediction accumulates with each subsequent multi-step-ahead prediction, making accurate predictions at longer forecasting horizons more difficult [39]. It is extremely important for this reason, that the base model is properly identified by selecting a near-optimal model that correctly explains the relationship between the explanatory and response variables.

Trial and error identification processes have been previously adopted for model selection [29,31]. Approaches involving selecting all (or significant numbers) of the potential predictors are unlikely to represent the best model because of the potential to include non-significant predictors; conversely, an insufficient number of model predictors might lead to poor performance in multi-step-ahead forecasts. Identifying a near-optimal model manually is therefore a difficult and time-consuming (and potentially impossible) task; and consequently, it is preferable to adopt an automated parameter selection process.

In the present work, the use of a simple automated model selection procedure designed for the calibration of ARX models is demonstrated to provide accurate one-step-ahead, i.e. hourly (1 h), predictions of the internal temperature evolution during a heatwave. A recursive strategy using sliding training and validation windows is adopted to provide hourly multi-step-ahead predictions for different forecasting horizons at: 3 h, 6 h, 12 h, 24 h, 48 h and 72 h periods. The forecasting accuracy and the 95% prediction intervals for the different forecasting horizons are evaluated using data from three different dwellings. The predictions across the various forecasting horizons were then repeated using ARMAX models in order to compare the forecasting accuracy and consistency of the results. The primary aim of this study is to assess the relative ability of ARX and ARMAX models to generate reliable multi-step-ahead temperature predictions.

2. Data selection

To stress test the predictive capabilities of a model in the context of ‘real-world’ overheating predictions, it is important that the model is tested and validated during a hot period in which external temperatures exceed those experienced during the previous (model training) period. According to the UK Met Office, based on the World Meteorological Organization definition, a heatwave is defined as a period of, “marked unusual hot weather (Max, Min and daily average) over a region persisting at least two consecutive days during the hot period of the year based on local climatological conditions, with thermal conditions recorded above given thresholds” [43,44]. For this purpose, three dwellings from the REFIT Smart Home dataset [34] were selected. The houses, all located in close proximity to the town of Loughborough in the English Midlands, experienced high temperatures but with markedly different temperature profiles, during the two-day heatwave of 30th June and 1st July 2015. During this short-duration extreme hot spell, the external air temperatures exceeded 30 °C in most regions of the UK [45]. The maximum external dry-bulb temperatures during that period set a new July record, with the highest temperature of 36.7 °C being observed at the Heathrow weather station [45]. On the hottest day: dwelling A (REFIT dwelling No. 12) exhibited a sudden indoor temperature spike exceeding 30 °C; dwelling B (REFIT dwelling No. 20) displayed a gradual increase in the internal temperatures with a lower peak of 27.6 °C, but with prolonged retention of elevated temperatures above 26 °C during the following night; and dwelling C (REFIT dwelling No. 7) displayed a sharp rise in temperature during the day but with a sudden drop in temperature overnight (Fig. 1).

The internal temperatures were logged at 30-min intervals in the upstairs bedrooms, to capture the most pronounced overheating. The weather data was recorded at the nearby Loughborough University weather station at 15-min intervals. As weather data and forecasts are not usually available in a sub-hourly resolution, and because hourly temperature data is more widely available, as a starting point for this work it was decided to down-sample the data by averaging the sub-hourly values to hourly mean values (centred on each hour). This reduces the number of time steps required to perform an extended forecasting horizon, and hence decreases the accumulation of errors due to the adopted recursive strategy in multi-step-ahead forecasts. However, it retains the ability to define the peak temperature reasonably accurately. Since the use of non-scaled data (i.e. non-normalised data) allowed more accurate predictions, the input data did not require transformation.

The data selected for the training and forecasting undertaken in this study extends across a five-week period from the 1st June 2015 to the 5th July 2015. During the first half of June, the external air temperatures \( T_{\text{air}} \) were considerably lower (ranging between 4 °C and 24 °C) than later in the month (Fig. 1). In the second half of June the external air temperatures showed a small increase in the daily mean, but with hourly peaks that did not exceed 24 °C until the 30th of June, when the external temperature suddenly increased to 31 °C. During the second day of the heatwave (1st July), the temperature rose even further and reached a peak of almost 36 °C. On the days following the short hot spell, the external daily temperature variations were very similar to those observed before the heatwave. The Global Horizontal Irradiance (GHI) showed similar daily variations before, during and after the heatwave (Fig. 1). Whereas the highest hourly GHI was recorded on the 7th June, when it peaked at almost 900 W/m², the peak on the 1st July was 750 W/m². Therefore, the higher external temperatures during the hot spell cannot be attributed to an increase in the solar irradiance.

For the whole of June, the internal temperatures \( T_{\text{int}} \) in the bedroom of dwelling B are consistently (1–3 °C) warmer than in the bedroom in dwelling A, which is 1–3 °C warmer than the bedroom of dwelling C (Fig. 1). The internal temperatures remain below 25 °C, and so would not be considered uncomfortable, until the 30th June when all the rooms reach approximately 26 °C. On the hottest day, 1st July, dwellings A and C heat up more noticeably than dwelling B, reaching maximal temperatures of 27.6 °C in dwelling B, 28.0 °C in dwelling C and 30.2 °C in dwelling A (Fig. 1).

3. Methods

3.1. Structure of the models

AutoRegressive models require that the input data used for the development of the model is stationary in order that the distribution of the observed and forecasted values is independent of time [38]. Hence, a time series can be considered stationary if the mean and variance of the data are constant [46] and if there are no significant trends and seasonal variations in the data [38]. To objectively determine if the data is

---

1 These are typically regression models which may omit relevant variables and/or include irrelevant variables.
stationary, unit root tests are adopted, with one of the most popular being the Augmented Dickey-Fuller (ADF) test [38]. The ADF unit root test was used to assess the stationarity of the input time series, with a probability value (p-value) threshold of 0.01. If the p-value of the ADF test is smaller than 0.01 (i.e. the ADF value is lower than the critical value for a specific sample size) the null hypothesis of a non-stationary time series can be rejected, and the alternative hypothesis of a stationary time series accepted.

Since analysis of all the input time series data used in this work satisfied the ADF unit root test (at the 99% confidence level) it can be concluded that the adopted data in this study is sufficiently stationary. As such, the input time series data does not require differentiation ($d = 0$) or further transformation to render it stationary. Without the use of the past residuals as predictors (i.e. with no Moving Average terms; $q = 0$) the model can therefore be denoted as an ARIMAX ($p, d = 0, q = 0, x$) model, or more simply described as an ARX ($p, x$) model.

To perform the forecasts at a specific time-step ($t$) and forecasting horizon ($h$), the model calibrates itself according to weightings applied to past internal temperatures ($T_{int}$) and to eXogenous inputs of past and/or forecasted weather data, consisting of $T_{ext}$ and GHI variables as recorded at the weather station. If the model is adopting the past residuals $q$ (i.e. the Moving Average order) in the forecasts ($q 
eq 0$), the model can be then denoted as an: ARIMAX ($p, d = 0, q, x$) model or more simply as an ARMAX ($p, q, x$) model.

The general equation of the model can be written in the form shown in equation (1).

$$T_{int}(t + h) = c + \phi_1 T_{int}(t + h - 1) + \ldots + \phi_n T_{int}(t + h - n) + \alpha_1 T_{ext}(t + h) + \ldots + \alpha_n T_{ext}(t + h - n) + \beta_1 GHI(t + h) + \ldots + \beta_n GHI(t + h - n) + \gamma_1 e(t + h - 1) + \ldots + \gamma_q e(t + h - q) + e(t + h)$$

Fig. 1. Hourly averages of the recorded internal temperatures ($T_{int}$) in dwellings A, B and C, and external air temperatures ($T_{ext}$) and Global Horizontal Irradiance (GHI) recorded at Loughborough University from the 1st June 2015 to the 5th July 2015.
\( \beta_n \) eXogenous coefficient (weight) of the past/forecasted GHI at lag \( n \) GHI \((t + h - n)\) observed or forecasted hourly Global Horizontal Irradiance at lag \( n \) before the forecasting horizon \( h \) [W/m²]

\( \gamma_n \) Moving Average coefficient (weight) of the past residual at lag \( q \) \( e(t + h - q) \) residuals: the hourly difference between the observed and forecasted internal temperatures at the time step \( t \) for the forecasting horizon \( h \) and lag \( q \) [°C]

\( q \) Moving Average order (i.e. the number of past residuals that are adopted to produce the forecasts)

\( e(t + h) \) forecasting error: the hourly difference between the forecasted and observed internal temperatures at the time step \( t \) for the forecasting horizon \( h \) [°C]

For one-step-ahead forecasts the model requires only the observed past internal temperatures \( (T_{\text{int}}) \) as AutoRegressive (AR) inputs, whilst for multi-step-ahead forecasts the model adopts partially (when \( 1 < h \leq n \)) or exclusively (when \( h > n \)) the observed internal temperature estimates (generated at previous time steps). Similarly, as eXogenous \( (X) \) inputs, the one-step-ahead forecasts require the observed past weather data \( (T_{\text{ext}} \text{ and GHI}) \) and the forecasted weather data for that specific time step \( (t + I) \). For multi-step-ahead forecasts, the model adopts partially (when \( 1 < h \leq n \)) or exclusively (when \( h > n \)) the forecasted weather data, which is assumed to be known with sufficient accuracy.

To perform the comparison of \( \text{ARX} \) and \( \text{ARMAX} \) models, the same AutoRegressive and eXogenous inputs identified for the \( \text{ARX} \) model were adopted for the \( \text{ARMAX} \) model. The moving average order, \( q \), was varied from 1 to 6 and the order producing the lowest Akaike Information Criterion (AIC) value for the specific dwelling was selected as the \( \text{ARMAX} \) model to be compared with the \( \text{ARX} \) model.

Since an extended training period of three weeks showed more consistent and accurate forecasts than either a 1 or 2-week training period, 21 days of data were used to train the regression coefficients \( (\phi_n, \alpha_n, \beta_n, \gamma_n) \) of the time series models. Hence, the training period extended from the 1st June at 00:00 to 21st June at 23:00, whilst the forecasting period started immediately after this, on the 22nd June at 00:00 (initial forecasting origin). For the purpose of this study the forecasts and their accuracy are analysed only during the one-week period of the heatwave event, from 28th June at 00:00 to 4th July at 23:00.

3.2. Model identification

In statistics, penalised likelihood criteria such as the AIC and the Bayesian Information Criterion (BIC) are often adopted for model selection [38]. Since the BIC measures the goodness of fit it is appropriate for explanatory models, whilst the AIC assesses the forecasting accuracy and is therefore better suited to predictive models [47]. The AIC (equation (2)) estimates the likelihood of the model to predict future values, which is penalised by the number of estimated parameters in the model (i.e. penalised likelihood). As such, the AIC addresses the trade-off between the goodness of fit of the model and the simplicity of the model. By automating the model calibration process, potentially viable models can be tested with all possible combinations of input variables.

The best model is then identified by selecting the combination of features (predictors) that result in the minimum value of the AIC estimator. According to Hyndman and Athanasopoulos [38], the model with the minimum value of the AIC is considered to be the optimal model for forecasting.

\[
\text{AIC} = -2 \ln(L) + 2N
\]

where:

\( \text{AIC} \) Akaike Information Criterion

\( L \) maximum likelihood of the estimated model

\( N \) number of estimated parameters in the model

In order to perform the model selection process in a reasonable amount of time (e.g. in less than 1 h) using code written in \( R \) [48] and using a single core (i.e. running the code in sequence), it was decided to limit the lag \( n \) (i.e. the number of previous time steps of data that are considered as predictors) to 5. This results in 131,072 possible model combinations resulting from the 17 available input parameters. The lagged inputs of \( T_{\text{int}}, T_{\text{ext}} \text{ and GHI} \) that resulted in the lowest AIC score with the \( \text{ARX} \) model were automatically selected. The selection process for the predictors was performed only once for each modelled zone during the training period (i.e. the first 21 days) and the selected model was then adopted to perform the rolling forecasts for that specific zone and dwelling. The number of AutoRegressive \( (p) \) and eXogenous \( (x) \) inputs chosen by the selection criteria for each model was automatically assigned to the names of the output files to enable model identification and facilitate cross-referencing of the extracted tables and plots.

3.3. Multi-step-ahead predictions

In ‘real-world’ applications a predictive overheating model would require forecasted weather data from one (or more) nearby meteorological station(s). Since the uncertainty of weather forecasts increases in proportion to the length of the forecasting horizon, their reliability several days ahead (particularly in a maritime climate) is questionable; as a result, using forecasting models to predict significantly long periods beyond the forecasting origin is likely to be unreliable. According to the UK Met Office, short-range (1–3 days ahead) weather forecasts are considered to be extremely accurate using data that is updated several times per day [49]. On the other hand, medium-range (3–10 days ahead) weather forecasts provide only a general picture of the weather on a day-to-day basis. For this reason, the developed models are constrained to forecasting \( T_{\text{int}} \) for the next 72 hourly time steps (3-day forecast) after the forecasting origin.

To create a multi-step forecast, the model performs a one-step-ahead forecast and then iteratively completes the multi-step-ahead forecasts for the next 72 h by adopting a recursive strategy. To achieve this, the model adopts a rolling forecasting origin (i.e. utilising sliding training and validation periods). This means that after each 72-h forecast, the model training window (21 days) moves forward by one time step (1 h), before recalibrating the regression coefficients (weights) of the previously selected predictors and then recalculating the forecasts. The model automatically stops when the forecasting window (of 72 h) reaches the end of the dataset. Once the rolling origin forecasts have been completed for the whole validation period, it is then possible to assess the forecasting accuracy.

3.4. Model validation

The accuracy of a forecasting model can only be evaluated based on how well it performs in relation to ‘new’ data [38], i.e. not how well the model fits the ‘past’ data which it was exposed to during the training period. In this study, the forecasting accuracy was evaluated only during the week of the heatwave (28th June at 00:00 to 4th July at 23:00) using scale-dependent error metrics: Mean Bias Error (MBE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The adjusted coefficient of determination \( (R^2_{\text{adj}}) \) was also calculated for reference. Whilst calculating \( R^2_{\text{adj}} \) during the training period (i.e. in-sample) can be useful in interpreting the goodness of fit between the model prediction and the measured data, it does not always indicate a good model for forecasting [38]. In fact, a good fit in the training period might be a sign of an over-fitted model; wherein the model matches the training data so closely that it loses the ability to generalise and forecast over the entire testing/validation period, with consequently poor forecasting performance. For these reasons, \( R^2_{\text{adj}} \) was used only to express the fit of the model over the validation (i.e. out-of-sample) period [38].
3.5. Reliability of forecasts

Knowing that a model is able to forecast accurately during a typical hot spell is not the only requisite characteristic of a reliable overheating forecasting model. Whilst sudden spikes of the internal and external temperatures can significantly decrease the short-term predictive accuracy of a model, it is important to consider that the main purpose of the model is to inform the occupants of the time and magnitude of impending overheating risks. In reality, it is likely that when faced with prolonged and/or severe overheating the occupants might take some mitigation actions (e.g. window opening, use of air conditioning etc.) and these sudden interventions could also significantly disrupt the forecasts, even when the model is slowly adapting to an overheating trend.

Although forecasts are often presented as deterministic point values, they can be better understood as the average value of a forecast probability distribution [38]. In real-world applications, occupants of a building may need to know in advance not only the likely future internal temperatures but also the reliability of such forecasts. Prediction intervals are commonly used to express the reliability of forecasts. They define the range of values within which a forecast is expected to lie with a specified probability. For a normal distribution, there is a 95% probability that the actual future temperature will lie within 1.96 standard deviations of the mean. Therefore, based on the central limit theorem, 1.96 standard deviations either side of the mean can be used as an estimate of the 95% prediction interval. Since the 95% prediction interval is slowly adapting as the time series evolves, and because it is produced based on past errors, it can be used to reliably inform the occupants of how reliable the forecast is expected to be for each forecasting horizon.

In order to produce the prediction interval, the standard deviation of the h-step forecast distribution (σf) has to be estimated for each forecasting horizon (h) [38]. This is done by calculating the standard deviation of the residuals (i.e. forecasting errors) at that specific forecasting horizon (h) over the preceding week of forecasts (with progressively shorter periods being subsequently adopted until the point where the first complete week of forecasted data is at least 5 days long). Once σf has been estimated it is possible to calculate the 95% prediction interval for each forecasting horizon (i.e. 1 h, 3 h, 6 h, 12 h, 24 h, 48 h and 72 h). The prediction interval (PIh) is iteratively recalculated at every time step as shown in equation (3).

\[
\text{PI}_h = T_{\text{fore}(t+h)} \pm k\sigma_f
\]

where:

- \( T_{\text{fore}(t+h)} \) Prediction Interval for the forecasting horizon h [°C]
- \( T_{\text{int}}(t+h) \) Forecasted hourly internal temperature at the time step t for the forecasting horizon h [°C]
- \( t \) Hourly time step [h]
- \( h \) Forecasting horizon (in hourly time steps) [h]
- \( k \) Coverage factor (k = 1.96 standard deviations for the 95% prediction interval)
- \( \sigma_f \) The standard deviation of the h-step forecast distribution

4. Results

4.1. Identified ARX models

The automatic selection procedure identified the ARX models with the following orders and predictors, as being optimal:

**Dwelling A:**
- Identified model: ARX (5, 6)
- AR inputs: \( T_{\text{int}}(t+h-1), T_{\text{int}}(t+h-2), T_{\text{int}}(t+h-3), T_{\text{int}}(t+h-4), T_{\text{int}}(t+h-5) \)

**eXogenous inputs:** \( T_{\text{ext}}(t+h), T_{\text{ext}}(t+h-4), GHI(t+h), GHI(t+h-1), GHI(t+h-2), GHI(t+h-4) \)

**Dwelling B:**
- Identified model: ARX (4, 5)
- AR inputs: \( T_{\text{int}}(t+h-1), T_{\text{int}}(t+h-2), T_{\text{int}}(t+h-3), T_{\text{int}}(t+h-4) \)
- eXogenous inputs: \( T_{\text{ext}}(t+h), T_{\text{ext}}(t+h-1), T_{\text{ext}}(t+h-2), T_{\text{ext}}(t+h-4) \), GHI(t+h)

**Dwelling C:**
- Identified model: ARX (3, 8)
- AR inputs: \( T_{\text{int}}(t+h-1), T_{\text{int}}(t+h-2), T_{\text{int}}(t+h-5) \)
- eXogenous inputs: \( T_{\text{ext}}(t+h), T_{\text{ext}}(t+h-2), T_{\text{ext}}(t+h-3), T_{\text{ext}}(t+h-5) \), GHI(t+h), GHI(t+h-1), GHI(t+h-4), GHI(t+h-5)

It can be observed (from the above descriptions) that the model for dwelling A has adopted more eXogenous predictors based on the previous time steps of the GHI, than \( T_{\text{ext}} \) for the model for dwelling B has used more terms based on the previous time steps of the \( T_{\text{ext}} \) than GHI; and the model for the dwelling C adopts considerably less AutoRegressive terms and a larger number of eXogenous inputs with an equal number of terms for both \( T_{\text{ext}} \) and GHI. It should be noted that there are also significant differences in the coefficient weightings of the various predictors. Overall, the AutoRegressive terms \( T_{\text{ext}} \) have the most dominant relative weights, whilst \( T_{\text{ext}} \) and GHI have only small and very small relative weights respectively. This means also that due to the lower relative weights of the eXogenous (weather) inputs, the models are globally less sensitive to the uncertainties associated with the external weather data.

4.2. Temperature forecasts

For dwelling A, the 1-h forecasts are very accurate and almost completely aligned with the observed values, with an \( R^2_{\text{MAE}} \) of 0.989. For the 3-h and 6-h forecasts, whilst the model is predicting accurately in relation to the peak temperature on the hottest day (1st July) (Fig. 2), there is a 2-h lag between the forecasted and observed peaks. For longer forecasting horizons (12–72 h), other than a delay of 1–2 h in predicting the timing of the peak temperature, the model under-predicts the peak internal temperature on 1st July, 28.4°C (12-h forecast) and 28.7°C (72-h forecast), compared to the measured peak of 30.2°C.

The model for dwelling A is also struggling to forecast the rapid drop in the internal temperatures on the afternoon of the 2nd July (from 26.2°C at 16:00 to 21.7°C at 21:00) at forecasting horizons of 3 or more hours. The sudden drop in temperature was caused by a rapid drop in the external temperature (Fig. 2), although it is possible occupants also opened windows to cool the room down. Overall, across the seven-day forecasting period, the model predicted with reasonable accuracy, with a maximum \( \text{MAE} \) of 0.69°C for the 72-h forecasts (Table 1).

For dwelling B, as for dwelling A, the 1-h forecasts are extremely accurate, with an \( R^2_{\text{MAE}} \) of 0.999. The 3-h, 6-h and 12-h forecasts are also reasonably accurate (Fig. 3). On the other hand, for longer forecasting horizons (24–72 h), the model tends to under-predict the peak temperature and struggles to accurately predict the retention of elevated temperatures between the 1st and 2nd July. Nonetheless, perhaps because dwelling B has a much smoother internal temperature profile (Fig. 3 cf. Fig. 2), the forecasts are more accurate than those for dwelling A for all the forecasting horizons as measured by the \( \text{MAE} \) and \( \text{RMSE} \) (Table 1). The tendency towards under prediction is evident in the \( \text{MBE} \). As for dwelling A, the \( \text{MBE} \) (in absolute terms), \( \text{MAE} \) and \( \text{RMSE} \) are all gradually increasing in magnitude as the forecasting horizon \( h \) increases.

\( ^2 \) Since is not possible to compare the coefficients for different variables directly, because they are measured on different scales (i.e. unstandardised coefficients), they are expressed as an average percentage weight for each specific input variable: \( T_{\text{int}}, T_{\text{ext}} \) and GHI.
Fig. 2. Dwelling A: observed, $T_{in}(t)$, and predicted, $T_{in}(t+h)$, hourly internal temperatures with hourly forecasting error, $e(t+h)$, and the 95% prediction intervals (grey bands) for 1 h, 3 h, 6 h, 12 h, 24 h and 72 h forecasting horizons, ARX model.
For dwelling C, despite the rapid fluctuations in the measured internal temperature, the model performed with reasonable accuracy throughout the entire week of the heatwave. The unusual temperature profile (i.e. a small increase followed by a sudden fall in the temperatures) on the 2nd July (Fig. 4) was difficult for the model to predict especially at longer forecasting horizons (12–72 h). Despite this challenge, the model performed with comparable accuracy to the model for dwelling A, and with bias and errors that gradually increase with extended forecasting horizons (Table 1).

### 4.3. Prediction intervals

For all three dwellings, the prediction intervals (grey bands in Figs. 2, 3 and 4) increase as the forecasting horizon (h) increases, with a notable increase from 3 to 6 h. As noted by Hyndman and Athanasopoulos [38], a common characteristic of prediction intervals is that they tend to gradually increase as the forecasting horizon lengths. The prediction interval also increases markedly after the heatwave, i.e. 2nd July for h ≥ 6 h, to about ± 1.5 °C and ± 1.25 °C for dwellings A and C respectively, and to approximately ± 0.75 °C for dwelling B.

With two brief exceptions, for all three dwellings and all forecasting horizons up to h = 12, the measured internal temperatures were within the prediction interval. The exceptions were dwellings A and C on 2nd July, where the indoor temperature showed a sudden dramatic decrease. For prediction horizons h = 24 and h = 72, the internal temperatures were not covered by the prediction interval at all times for any of the dwellings on the 1st July and, more notably on the 2nd July. The observed temperature was above the prediction interval for dwellings B and C, and over then under for dwelling A.

When forecasted temperatures lie outside the prediction interval for a prolonged period it suggests that the model is not sufficiently reliable or that the response of the room to changes in ambient conditions differs from that which occurred during the training period. These matters are discussed further below.

### 4.4. Comparison of ARX and ARMAX models

For all three dwellings, ARMAX models (q ≠ 0) were developed using the same AutoRegressive and eXogenous terms as in the ARX models (see section 4.1). By varying q, the lowest AIC values were determined as q = 5, q = 4 and q = 6 for dwellings A, B and C respectively. Whilst this would suggest that, at least in theory, the ARMAX models would provide better forecasting accuracy than ARX models, other aspects need to be considered. Firstly, the AIC values are determined from the training period for one-step-ahead forecasts only; and secondly, when making actual forecasts, the future residuals cannot be computed a posteriori (and therefore estimates cannot be obtained), and consequently they are set to zero. This means that whilst for one-step-ahead forecasts (h = 1) the model is using q residuals in the calculation, when 1 < h ≤ q the moving average inputs gradually include more zero values, and all are null once h > q.

Comparison of the predictive accuracy metrics using the ARMAX model (Table 2) with the corresponding accuracy metrics using the ARX model (Table 1) indicates that, with proper identification of the ARX model, the ARMAX models provide little if any overall improvement. For dwelling A, the ARMAX model yielded a lower MBE but at the expense of an increased RMSE value, for dwelling B, the MBE and MAE were very similar but the RMSE was slightly worse, and for dwelling C, MBE, MAE and RMSE were all worse for the ARMAX model for all of the forecasting horizons.

### 5. Discussion

The aim of this work is to lay the theoretical foundation for an in-home device that could provide an early warning of likely elevated temperatures. Model automation is an extremely important feature of such a device since it obviates the need for manual intervention, trial and error procedures, or model identification by an expert. In principle, therefore, it might be possible to develop a device that needs only a sensor to record the internal zonal air temperature and an internet (or cellular mobile) connection to continuously access and download the weather forecast for a specific location. After an initial training period, the device would be able to automatically select an appropriate model for the specific room before continuing to perform ongoing forecasts of the internal temperatures.

Interestingly, the parameter weightings of the derived models suggest that they are relatively immune to the uncertainty in the input weather data. Therefore, even if the derived models were to rely upon forecasted weather data from more distant meteorological stations or on interpolated data (with assumptions about topographical and microclimate effects), the predictive accuracy may not degrade, which is a useful attribute if the device were deployed in remote locations.

The work undertaken here, concurs with the findings of Mustafaraj et al. [30] in that the ARX and ARMAX models produced similar results, however, as Ríos-Moreno et al. [29] also observed, the ARX models were generally a little better overall such that the additional complexity of using an ARMAX model does not appear to be justified. The accuracy of the ARX predictions in the rooms that responded more dramatically to external temperature changes (dwellings A and C), are poorer than those reported previously. Ferracuti et al. [33], quoted an RMSE of 0.33 °C for 3-h summertime forecasts in buildings using ARX models, whilst here the values ranged from 0.14 °C in dwelling B, to 0.51 °C and 0.62 °C respectively in dwellings A and C. Likewise, the 2-h summertime forecasts using ARMAX models reported by Mustafaraj et al. [22], produced a MAE range of 0.11–0.19 °C, which is better than the 3-h forecasts for dwellings A and C of 0.35 °C and 0.44 °C, respectively. The results are however, better than those of Ashtiani et al. [32], whose time series regression model was not able to forecast accurately during a heatwave event.

There are several reasons for these differences: firstly, the studies of Mustafaraj et al. [30] and Ferracuti et al. [33] were performed in office buildings with extensive measured data used for the development of the forecasting models. The internal temperatures of offices tend to be less affected by ambient conditions than those in dwellings and individual

---

**Table 1**

Forecasting accuracy for the week of the 2015 heatwave, ARX models.

<table>
<thead>
<tr>
<th>Forecasting horizon h (hours)</th>
<th>Dwelling A</th>
<th>Dwelling B</th>
<th>Dwelling C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2_{fl}$ (0–1)</td>
<td>MAE (°C)</td>
<td>MAE (°C)</td>
</tr>
<tr>
<td>1</td>
<td>0.989</td>
<td>-0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>0.955</td>
<td>-0.08</td>
<td>0.35</td>
</tr>
<tr>
<td>6</td>
<td>0.921</td>
<td>-0.16</td>
<td>0.53</td>
</tr>
<tr>
<td>12</td>
<td>0.910</td>
<td>-0.24</td>
<td>0.57</td>
</tr>
<tr>
<td>24</td>
<td>0.898</td>
<td>-0.35</td>
<td>0.57</td>
</tr>
<tr>
<td>48</td>
<td>0.887</td>
<td>-0.45</td>
<td>0.62</td>
</tr>
<tr>
<td>72</td>
<td>0.876</td>
<td>-0.56</td>
<td>0.69</td>
</tr>
</tbody>
</table>
occupants have less, and sometimes no, personal control over the internal environment. Secondly, the previous work was undertaken during mild summer days with no sudden temperature spikes and peak indoor temperatures of approximately 24 °C [30] to 26 °C [33]. However, when Mustafaraj et al. used their ARX model in a second study [31], in which the internal office temperatures were higher (28–29 °C)

Fig. 3. Dwelling B: observed, $T_{in}(t)$, and predicted, $T_{in}(t + h)$, hourly internal temperatures with hourly forecasting error, $e(t + h)$, and the 95% prediction intervals (grey bands) for 1 h, 3 h, 6 h, 12 h, 24 h and 72 h forecasting horizons, ARX model.
Fig. 4. Dwelling C: observed, $T_{int}(t)$, and predicted, $T_{int}(t + h)$, hourly internal temperatures with hourly forecasting error, $e(t + h)$, and the 95% prediction intervals (grey bands) for 1 h, 3 h, 6 h, 12 h, 24 h and 72 h forecasting horizons, ARX model.
even though the diurnal range of indoor temperature was similar throughout the week, the MAE of the 2-h predictions increased to 0.37–0.49 °C, which is slightly worse than the 0.12–0.44 °C achieved in this study for 3-h forecasts during a heatwave. Thirdly, in the office studies of both Mustafaraj et al. and Ferracuti et al. many more inputs were provided to the models than in the work reported here.3 The difficulty that the ARX and ARMAX models, deployed in this study, had in making predictions during abnormal temperature events (and over longer forecasting horizons) is not surprising. Firstly, the models can only be trained based on past events, so the prediction for sudden, rare and more extreme events will always be difficult. Secondly, during such events, the occupants of homes may behave differently; abnormally even. Mitigating actions during a heatwave could include, opening windows and even doors, closing the curtains during the day, turning on portable fans or even using portable air conditioning units. Models learn slowly and so whilst such actions will be incorporated in the model the quality of immediate forward predictions, even for only 3 h ahead will be degraded. Whereas mitigation actions at longer forecasting horizons (3–72 h) might generate occasional false positives (i.e. lower temperatures occurred than were predicted, due to the intervening actions taken by the occupants) access to early information regarding the expected room temperatures allows the occupants to take preventative actions. As such, the forecasts can be viewed as a prediction of what will happen if no one intervenes (beyond the established patterns of operation). From a health perspective, this is useful information since it allows the occupants (or their carers) to take action to lower the indoor temperatures in order to contain them within a comfort range or within acceptable heat-stress levels. Because at shorter forecasting horizons (1–3 h) the predictions are considerably more accurate and consistent, and because the target ambulance response times in the UK for non-urgent calls are within 3 h for 90% of the calls [50], a warning device could be set to dispatch the emergency services in advance in order to reach vulnerable occupants (i.e. those most at risk during hot weather) in a timely manner. Additional sensors, for example to detect window opening or internal air velocities, could assist the model, but this adds cost and complexity and only deals with one of the many possible occupant responses.

Future work will focus initially on further improvements to the modelling procedure and understanding the factors that affect the models’ predictive accuracy. One approach that will be explored is the use of non-linear models (e.g. ANN models etc.), which were shown in the study by Mustafaraj et al. [31] to produce more accurate forecasts than linear ARX models. In addition, the models will be tested on datasets that contain information on window opening patterns in order to examine their effect on the predictions of the models.

This work has examined three, specifically-selected rooms, in different dwellings located in the same town, with hourly temperatures recorded over just one summer period. Future endeavours will entail testing the modelling process and quantifying the models’ accuracy, for many more rooms, households, dwelling types and locations.

Ultimately, it is hoped that forecasts of sufficient reliability could be provided to vulnerable occupants (and their carers) several days in advance (24–72 h) with minimal monitoring and at a low cost. This would allow occupants, carers and perhaps the emergency services adequate time to prepare for an impending response. Whilst the very reliable short-term forecasts (1–12 h) would allow the targeted deployment and triaging of emergency services.

6. Conclusions

The potential for statistical models to predict indoor temperatures during heatwaves has been investigated using hourly data from three bedrooms, in three houses, located close to the town of Loughborough in the UK Midlands. During the monitoring period, there was a two-day heatwave during which the external dry-bulb temperature exceeded 35 °C. The AIC was adopted to automatically identify a near optimal forecasting model, specific to each room, using data from before the period of hot weather. Recursive multi-step-ahead forecasts were made by both ARX and ARMAX models using a rolling forecasting origin. These provided predictions for forecasting horizons of 1, 3, 6, 12, 24 and 72 h for the whole week of the heatwave. The accuracy of the predictions over that week were evaluated using the MBE and $R^2_{adj}$ as measures of the bias and out-of-sample fit of the models, and MAE and RMSE to assess the forecasting accuracy of the models. The 95% prediction intervals were computed for the heatwave week to express the reliability of the forecasts at different forecasting horizons.

Comparison between the ARX and ARMAX models showed that whilst they produce almost identical one-step-ahead forecasts when longer multi-step-ahead forecasts are performed, with a recursive strategy, the ARX models were simpler to derive and offered slightly more consistent, reliable and accurate predictions. The ARX models produced an MAE below 0.7 °C during a heatwave week for all three dwellings and for all of the forecasting horizons up to 72 h. The internal temperatures tended to be under-predicted for two dwellings, MBE up to ~0.56 °C, but over-predicted for the other, MBE, 0.31 °C. The range of the 95% prediction interval varied from ±0.75 °C in one dwelling to ±1.50 °C in the dwellings that responded more dramatically to the elevated temperatures during the heatwave. With very limited local exceptions, the actual temperatures were within the prediction interval for all forecasting horizons up to 12 h.

Overall the early findings of the work reported here suggest that highly detailed building information is not required to produce reasonable forecasts of indoor temperatures in free-running (i.e. without mechanical cooling and heating) dwellings.

This points to the potential for using time series forecasting as part of an overheating early-warning system in buildings, especially those housing vulnerable occupants or contents. Future work will explore alternative non-linear modelling approaches and examine the effect of

| Table 2: Forecasting accuracy for the week of the 2015 heatwave, ARMAX models. |
|-----------------|-----------------|-----------------|
| Forecasting horizon h (hours) | Dwelling A | Dwelling B | Dwelling C |
| $R^2_{adj}$ (0–1) | MBE (°C) | MAE (°C) | RMSE (°C) | $R^2_{adj}$ (0–1) | MBE (°C) | MAE (°C) | RMSE (°C) | $R^2_{adj}$ (0–1) | MBE (°C) | MAE (°C) | RMSE (°C) |
| 1 | 0.989 | 0.00 | 0.17 | 0.26 | 0.999 | -0.01 | 0.04 | 0.05 | 0.989 | -0.02 | 0.12 | 0.21 |
| 3 | 0.865 | 0.01 | 0.56 | 0.75 | 0.977 | -0.04 | 0.14 | 0.16 | 0.932 | -0.14 | 0.43 | 0.62 |
| 6 | 0.797 | 0.00 | 0.70 | 0.94 | 0.932 | -0.10 | 0.24 | 0.28 | 0.786 | -0.40 | 0.80 | 1.05 |
| 12 | 0.786 | 0.00 | 0.70 | 0.96 | 0.898 | -0.15 | 0.30 | 0.36 | 0.865 | -0.36 | 0.67 | 0.86 |
| 24 | 0.774 | 0.11 | 0.67 | 0.98 | 0.842 | -0.25 | 0.31 | 0.45 | 0.887 | -0.45 | 0.64 | 0.78 |
| 48 | 0.831 | 0.20 | 0.61 | 0.85 | 0.774 | -0.38 | 0.42 | 0.53 | 0.865 | -0.58 | 0.71 | 0.85 |
| 72 | 0.842 | 0.23 | 0.59 | 0.81 | 0.684 | -0.45 | 0.51 | 0.62 | 0.831 | -0.70 | 0.79 | 0.94 |

[3] Mustafaraj et al. [30,31]: internal and external temperatures; internal and external relative humidity; supply air flow-rate, air temperature and relative humidity of the air handling units; chilled water temperature of the chillers; hot water flow temperature to the fan coil unit. Ferracuti et al. [33]: internal and external temperatures, solar gains, internal gains and thermal gains.
windows opening and other occupant interventions on the predictive accuracy of the models. Finally, testing of the prototype forecasting models on larger datasets will be carried out in order to quantify the reliability of predictions for different rooms, dwelling and household configurations across a wide range of geographic locations.

Declarations of interest

None.

Acknowledgements

This research was made possible by Engineering and Physical Sciences Research Council (EPSRC) support for the London-Loughborough (LoLo) Centre for Doctoral Training in Energy Demand (grant EP/L01517X/1). Monitored data, indispensable to this study, was made available by the open access REFIT Smart Home dataset [34], which was funded by the EPSRC. ‘REFIT: Personalised Retrofit Decision Support Tools for UK Homes using Smart Home Technology’ (grant EP/K002457/1).

Nomenclature

$\alpha_n$ eXogenous coefficient (weight) of the past/forecasted external air temperature ($T_{ext}$) at lag $n$

$\text{AIC}$ Akaike Information Criterion

$\text{ADF}$ test Augmented Dickey-Fuller test

$\text{ARIMAX (p, d, q, x)}$ AutoRegressive Integrated Moving Average with eXogenous inputs

$\text{ARMAX (p, d, q)}$ AutoRegressive Moving Average time series with eXogenous inputs

$\text{ANN}$ Artificial Neural Network

$\text{ARX (p, x)}$ AutoRegressive time series with eXogenous inputs ($d = 0$; $q = 0$)

$\text{BIC}$ Bayesian Information Criterion

$\beta_n$ eXogenous coefficient (weight) of the past/forecasted Global Horizontal Irradiance (GHI) at lag $n$

$c$ intercept (regression constant) [°C]

$\gamma_q$ Moving Average coefficient (weight) of the past residual at lag $q$

$d$ integration order: adopted order of differencing required to make the input data stationary

$e (t + h - g)$ residuals: the hourly difference between the observed and forecasted internal temperatures at the time step $t$ for the forecasting horizon $h$ and lag $q$ [°C]

$e (t + h)$ forecasting error: the hourly difference between the forecasted and observed internal temperatures at the time step $t$ for the forecasting horizon $h$ [°C]

$\text{GHI}$ observed or forecasted hourly Global Horizontal Irradiance [W/m²]

$h$ forecasting horizon (h-step forecast) [h]

$k$ coverage factor ($k = 1.96$ standard deviations for the 95% prediction interval)

$L$ maximum likelihood of the estimated model

$\text{MAE}$ Mean Absolute Error [°C]

$\text{MBE}$ Mean Bias Error [°C]

$n$ lag ($n^{th}$ previous time step of the input variable) [h]

$N$ number of estimated parameters in the model

$p$ AutoRegressive inputs: number of past observed values considered as predictors

$\phi_n$ AutoRegressive coefficient (weight) of the past internal Temperature ($T_{int}$) at lag $n$

$P_{th}$ 95% prediction interval for the forecasting horizon $h$ [°C]

$q$ Moving Average (MA) order: number of past residuals considered as predictors

$R^2_{adj}$ adjusted coefficient of determination [0–1]

RMSE Root Mean Squared Error [°C]

$t$ the standard deviation of the h-step forecast distribution

$T_{ext}$ observed or forecasted hourly external air Temperature [°C]

$T_{int}(t)$ observed hourly internal Temperature at the time step $t$ [°C]

$T_{int}(t + h)$ forecasted hourly internal Temperature at the time step $t$ for the forecasting horizon $h$ [°C]

$x$ eXogenous inputs: number of external variables adopted as predictors

References


