INTERNATIONAL BUILDING PERFORMANCE SIMULATION ASSOCIATION ENGLAND

Proceedings of BSO-V 2020
5th Building Simulation and Optimization Virtual Conference
21-22 September 2020
Loughborough, UK

One-Step Ahead Modelling of building thermal dynamics: A Comparison of Backward Selection and LASSO approaches

Argyris Oraiopoulos[1], Bianca Howard[1]
[1]Building Energy Research Group (BERG)
School of Architecture, Building and Civil Engineering (ABCE)
Loughborough University
Loughborough, UK

Abstract

The computation time and the infinite possibilities in model structures of data driven models often hinder the efficient development of accurate models. This paper presents a systematic approach for selecting the appropriate model when forecasting the temperature dynamics in non-residential buildings, using different streams of data. The main objective of the work is to evaluate the presented approach by comparing the results to those obtained by a typical backward elimination method.

The workflow delivers the selection of the appropriate features in order to represent the system accurately in a parsimonious model, by setting the initial model structure search space and then estimating the parameters, using the least absolute shrinkage and selection operator (LASSO) procedure. The analysis is performed on a case study educational building complex at Loughborough University in the Midlands, UK. The input data comprise of multiple features including internal room air temperatures, external air temperatures, and HVAC related data such as valve positions and fan speeds, measured sub-hourly over a winter period between 2018 and 2019.

The results confirm that the specified automated workflow enables accurate estimates of indoor air temperature using considerably less computational effort than the backward selection approach. However, the final form of the models identified could lead to poor control performance.

Introduction

Constructing models from empirical data has been a fundamental element in science throughout history. However, advances in monitoring equipment of building energy systems in recent years, have allowed for data driven modelling methods to be explored in much greater depth than ever before in the field of building physics. Depending on the available data and the required output, various models have been applied for achieving objectives, such as energy demand reduction, fault detection and optimisation of HVAC systems operation. However, the large plethora of modelling techniques that exists is not fully explored often, mainly due to time restrictions or lack of experience.

The identification of the correct model can be a challenging task that is often referred to as system identification, a term whose origins can be traced to Zadeh (1956) for the model estimation problem for dynamic systems. Ljung (2010) defined system identification as the "*art and science of building mathematical models of dynamic systems from observed input–output data*" (Ljung (2010)). Selecting the optimal input data in order to represent the system accurately in a parsimonious model, is a process that requires in depth exploration of all the input variables and has been termed feature selection. According to Liu and Hiroshi (1998), feature selection is the process of eliminating features from the database that are irrelevant to the task to be performed. Guyon and Elisseeff (2003) suggested that there are three main objectives in feature selection: firstly improving the prediction performance of the models, secondly producing faster and more cost-effective models, and thirdly providing a better understanding of the underlying processes in a given system.

Feature selection methods have been categorised in filter, wrapper and embedded (Chandrashekar and Sahin (2014)). Filter methods select features by ranking them based on correlation criteria regardless of the model, while wrapper and embedded methods select the optimal state (e.g. a subset of features) based on the performance of the model, therefore consistently delivering higher performance (Jović et al. (2015)). All methods consist of a search strategy and a rating methodology. The search strategy describes the selection of specific states (e.g. a subset of features) for rating and the rating methodology determines how those states are rated (Rätz et al. (2019)).

Wrapper methods have been used extensively in developing models for building energy forecasting. Zhang and Wen (2019) developed a systematic feature selection procedure that included a wrapper method to determine the best feature set in developing a data-driven building energy forecasting model. The proposed method outperformed that built with a filter method alone in two case study buildings. Rätz

et al. (2019) used wrapper methods to develop an efficient methodology for exploring the potential of data-driven machine learning models in modelling building energy systems. A common wrapper method is the stepwise regression method that has two main classes, the forward selection and the backward elimination (Draper and Smith (2000)). The forward selection involves the stepwise addition of features, starting with no variables, until the model is not statistically significant improved, while the backward elimination involves starting with all the nominated variables and stepwise deleting features until the statistically significant deterioration of the model fit does not allow for any further elimination. Vu et al. (2015) used backward elimination processes to select the most appropriate variables and develop a multiple regression model for monthly forecasting of electricity demand in the state of New South Wales in Australia. Amiri et al. (2015) also used backward elimination to reduce the number of parameters in their model and only include the most effective parameters in predicting energy consumption in the USA. Geneidy and Howard (2020) used the backward elimination is selecting the features of a model that was later applied to identify the factors affecting the contracted energy flexibility potential of homes.

Embedded methods have also been popular in developing models for predicting energy demand in buildings. Fan et al. (2014) applied recursive feature elimination to select the optimal inputs for next-day energy consumption and peak power demand predictions of a case study building in Hong Kong. Candanedo et al. (2017) also used recursive feature elimination for selecting the variables in the prediction of energy demand in a low energy house in Belgium.

An increasingly popular embedded method is the least absolute shrinkage and selection operator (LASSO). This approach minimizes the residual sum of squares, with a constraint of keeping the sum of the absolute value of the coefficients less than a constant. This constraint results in producing some coefficients that are exactly zero and hence gives parsimonious models (Tibshirani (1996)). Ma and Cheng (2016) used the lasso technique for identifying the influential features on the regional energy use intensity of residential buildings in New York. Wang et al. (2019) applied lasso to improve the accuracy and reliability of the occupancy detection of their model. Chen et al. (2018) proposed lasso in selecting the most influential variables to establish models for predicting building energy consumption. Finally, Suryanarayana et al. (2018) chose lasso to perform feature selection in developing a model for operational day-ahead heat demand forecasting in district heating systems.

Whilst these methods have been used for more general forecasting of building electricity demand, their use for detailed building dynamics models for use for control has been limited.

This paper will test the effectiveness of two popular methods,the backward elimination and lasso as feature selection methods, for the modelling of the internal temperature in a case study educational building, in UK. The aim is to identify efficient methods of identifying variables that enable the control mechanisms of building thermal dynamics.

Methodology

This study presents a comparison of feature selection methods in modelling the internal temperature of a teaching space, in an educational building. The following subsections describe the data used as the basis for modelling, the Ordinary Least Squares (OLS) and LASSO approaches used, and the validation techniques and metrics.

Data

The work used data from a case study educational building complex at Loughborough University, located in the Midlands, UK. The building, situated in the west park of the campus, was refurbished in 2016 and it comprises of a central learning and exhibition zone, lecture theatres, seminar rooms and an informal learning area.

Data were acquired from the building management system for the winter period 2018-2019. After a first screening process to identify missing data, the internal temperature of a lecture theatre (Teaching Room 1) was obtained with continuous data for the period between 17 January 2019 and 15 February 2019, in 15 minutes intervals. This is presented in Figure 1.

A number of variables that were considered to have an impact on the formation of the internal temperature in Teaching Room 1 and play a key role in controlling the indoor thermal environment, were also obtained, in the same resolution (15 minutes intervals) and for the same period (between 17 January 2019 and 15 February 2019). These are variables related to operation of the air handling unit that serves Teaching Room 1 as well as the temperature of the adjacent exhibition space and the external temperature, which is continuously measured by the facilities management team, using high accuracy probes. In summary, the following variables were considered for the modelling:

- Internal air temperature
- Cooling valve demand
- Heating valve demand
- CO_2 fan speed demand
- Damper calculated demand
- Outside air temperature

The cooling valve demand, heating valve demand, the CO_2 fan speed and damper calculated demand variables were considered as they were the main control signals that define the operation of the air handling unit. The internal temperature is the variable of interest for prediction and the outside air temperature

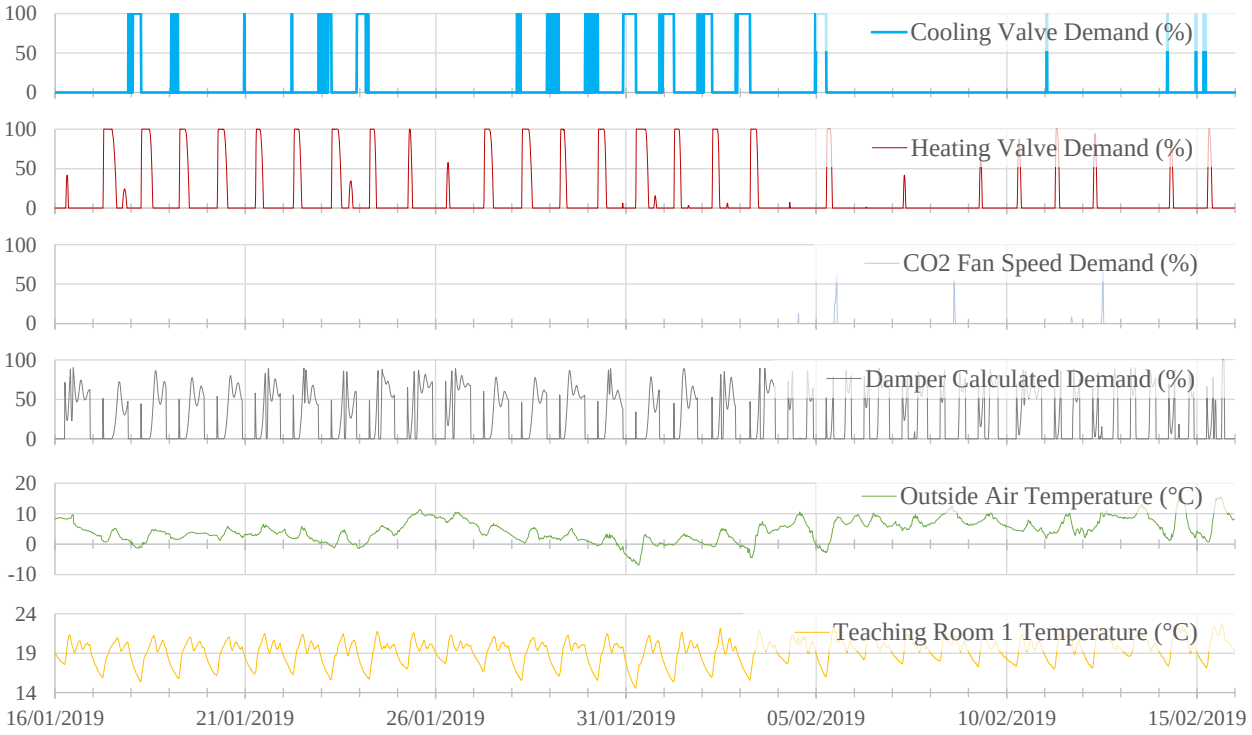

Figure 1: Data

is a common external factor that affects building indoor temperature. The values of these parameters over the period of evaluation is shown in Figure 1. These variables formed the basis for the modelling approach and the analysis presented in this paper.

Modelling

The proposed model is an Ordinary Least Squares (OLS) calculation of the internal temperature, based on past values of both endogenous and exogenous variables, for the calculation of one time step ahead forecasts. Endogenous variables are the own lags of the dependent variable, in this case of the internal temperature of Teaching Space 1. Exogenous variables are all the rest of the inputs to the model, other than the own lags of the internal temperature of Teaching Space 1. The general form of the model is given by the following equation:

$$y_{t+1} = \sum_{i=0}^{T_n} \beta_i y_{t-i} + \sum_{j=1}^{X} \sum_{i=0}^{T_x} \delta_{j,i} x_{j,t-i} \qquad (1)$$

Where y_t denotes the internal temperature of Teaching Room 1 at time t, $x_{j,t}$ denotes the exogenous variable j at time t, i is the number of own lags of the endogenous or exogenous variables,$beta_i$ is the parameter relating lag i of the endogenous variable to its value at $t+1$, $\delta_{j,i}$ is the parameter relating lag i of the exogenous variable x_j to the endogenous variable at time $t+1$, X defines the number of exogenous variables, T_n defines the number of lags of the endogenous variable to be explored and T_x defines the number of lags of the exogenous variables to be explored. .

The main aim of this model is to provide parsimony and improve the understanding of the dynamics in the building's thermal behaviour. Its composition allows for the future control of variables that are part of the HVAC system in order to achieve the desired thermal conditions in a space and allow energy efficiency and energy flexibility approaches to be implemented. The simplicity of the OLS calculation further allows the exploration of various feature selection methods.

Feature Selection

This paper explores two different feature selection techniques, the backward elimination and the least absolute shrinkage and selection operator (LASSO). These will be evaluated against the base case of developing a model with no feature selection technique applied, but they will also be compared against each other in terms of model accuracy and parsimony.

The backward elimination feature selection involved initially selecting a significance level $p < 0.05$, then the model was first fitted with all the nominated variables and the significance level of each variable was calculated. A stepwise process followed where the variable with the highest p-value was identified and eliminated if its p-value was greater than 0.05. The model was then fitted again with the remaining variables. This process was repeated until all variables left in model were significant with a p-value less than 0.05. This threshold is consistent with Vu et al. (2015) and also suggested in literature Montgomery (2011). Figure 2 illustrates the backward elimination

feature selection process.
The LASSO technique fits the model with a regularisation technique. This means the feature selection was applied by penalising the magnitude of coefficients of features together with minimising the error between measured and predicted values on the internal temperature (residual sum of squares or "RSS"). More particularly, LASSO performed an L1 regularisation, by adding a factor of sum of absolute value of coefficients ("SAC") in the objective function (which is minimising the error). This factor is depicted by the α (alpha) parameter, which balances the amount of weight given to minimising the residual sum of squares compared to minimising the sum of absolute values of the features' coefficients. The LASSO model was implemented using the statsmodels package where the objective function to be minimised is the following:

$$Objective = \frac{0.5 * RSS}{n} + \alpha * SAC \qquad (2)$$

It is clear that if the value of α is zero then the LASSO performs a normal OLS without feature selection. For values of $\alpha > 0$ the LASSO should result in producing some coefficients that are equal to zero and hence reduced model complexity, allowing parsimonious models to be developed. Figure 3 illustrates the LASSO feature selection process.

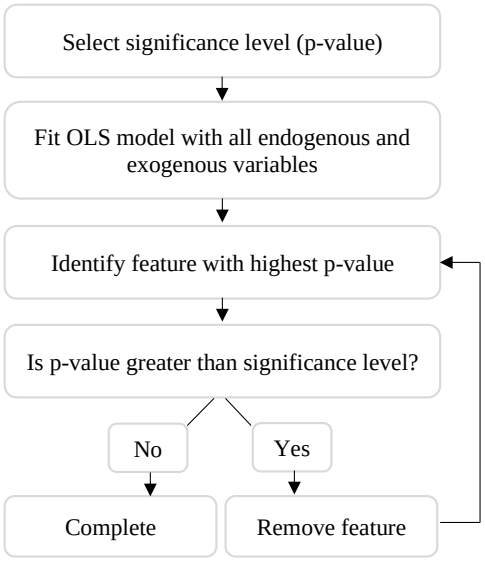

Figure 2: Backward elimination feature selection process

The nominated variables that form the features of the model are the own lags of the endogenous and the own lags of the exogenous variables. For the endogenous variable this work has selected the value at time t as well as the lags of up to 16 time steps (T_n), for the one-time-step ahead prediction $(t+1)$. With a single time step being the 15 minutes interval of the data resolution, 16 time steps are equal to 4 hours worth

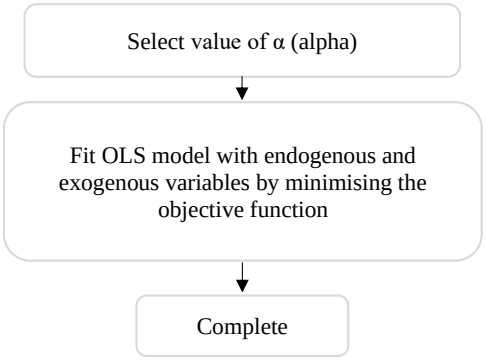

Figure 3: LASSO feature selection process

of data. For the exogenous variables in this work, the values at time t and those at the lags of up to 4 time steps (T_x) (1 hour worth of data) were considered.
For the computation of the feature selection, the coding language Python (3.7.1) was used and in particular the package statsmodels, installed on a standard laptop (i5 Processor @1.60GHz, 8GB RAM, 64-bit operating Windows 10 system).

Validation

The presented work has been validated using internal split validation. The data were split into training and testing sets. The training dataset was from 12:00am on 17 January 2019 until 11:45pm on 4 February 2019 (about 70% for the dataset) and the testing dataset was from 12:00am on 5 February 2019 until 11:45pm on 15 February 2019 (about 30% of the dataset). The training dataset was used to fit the model and calculate the coefficients of all the selected features and the testing dataset was used to calculate the one time step ahead $(t+1)$ predictions at each time step (t). The measures of error used to evaluate the results, given below for n observations, were based on the residuals between the measured value at $(t+1)$, marked with y in the equations below, and the prediction (the simulated) marked with $\hat{y}$.

- Coefficient of determination (R^2) is the proportion of the variance in the dependent variable that is predictable from the independent variables.

$$R^2 = \frac{n \times (\sum_{i=1}^{n} y_i \times \hat{y}_i) - \sum_{i=1}^{n} y_i \times \sum_{i=1}^{n} \hat{y}_i}{\sqrt{(n \times \sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2) \times (n \times \sum_{i=1}^{n} \hat{y}_i^2 - (\sum_{i=1}^{n} \hat{y}_i)^2)}}$$

- Root Mean square Error (RMSE): the standard deviation of the prediction errors.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}$$

- Coefficient of Variation of RMSE (CVRMSE): the normalised RMSE by the mean value of the measurements.

$$CVRMSE = \frac{1}{\bar{y}} \times \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \times 100$$

- Mean Absolute Error (MAE): the mean of the absolute difference between the measured and the simulated values.

  $MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}$

- Mean Absolute Percentage Error (MAPE): the average of absolute percentage errors .

  $MAPE = \frac{\sum_{i=1}^{n} |\frac{y_i - \hat{y}_i}{y_i}|}{n} \times 100$

- Mean Biased Error (MBE): the mean of the difference between the measured and the simulated values

  $MBE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n}$

- Normalised MBE (NMBE): the normalised value of the mean biased error that allows for comparable results.

  $NMBE = \frac{1}{\bar{y}} \times \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \times 100$

These measures of error have been extensively used in the field of forecasting and energy demand studies. Each one depicts a different aspect of the prediction as it will be discussed when presenting the results.

Results

The results from three models for the prediction of the internal temperature of Teaching Room 1 are presented here:

- the OLS with no feature selection
- the OLS BE with backward elimination feature selection
- the OLS LASSO (with three different α values: 0.1, 0.5 and 1.0)

The comparison of the one-time-step ahead predictions of all three models to the measured testing data is presented in Figure 4 and shows a good overall agreement, with the OLS LASSO model underpredicting some of the daily peaks of the internal temperature.

The measures of error for the one-time-step ahead predictions in Table 1 show that the performance of all models is high, with the OLS with backward elimination model performing almost identical to the OLS model with all the nominated features selected. The values of R^2 are very close to 1, which was to be expected after inspecting Figure 4. The measures RMSE, MAE and MBE indicate that the magnitude of the error for the test dataset is low. The MAE shows that on average, the error is below 0.2°C, while the RMSE gives more emphasis on larger differences, skewing the average difference to 0.1-0.2°C for all models. The MBE, which is the only measure that can capture the over and under prediction of the models, shows that, on average, all models under-predict the measured values, with an average error value of less than 0.1°, which is a relatively small value.

The measures CVRMSE, MAPE, and NMBE, allow for a relative comparison showing the error as a percentage of the mean value for every model. The value of CVRMSE indicates that the RMSE is 0.6% of the mean value of the measured series in the OLS and OLS with backward elimination and between 1.07-1.17% in the OLS with LASSO model, values that are considerably small. The MAPE shows that in terms of absolute error, the mean is proportional to less than 0.5% of the mean measured value in the cases of OLS and OLS with backward elimination and less than 1% in the case of OLS with LASSO. The NMBE, indicates that the average errors are up to 0.7% of the mean value in the case of the OLS LASSO model, while less than 0.1% in the other two models.

It is clear that as the value of α increases, the performance of the model degrades. This was to be expected since, the higher values of α, the less features are selected in the model, affecting its performance. Table 2 shows the features that were selected for each model and the calculated coefficients. It can be observed that the OLS with backward elimination model has less than half of the features (19 features) compared to the initial OLS model (42 features).

The OLS LASSO model has a considerably different structure than the OLS with backward elimination, yet the number of features selected is comparable, especially for α values of 0.5 (21 features selected) and 1.0 (20 features selected). The value of the internal temperature at the value at time t, for the prediction at $t + 1$, is the one with the most influence in all models. This was to be expected as the values of internal temperature in high resolution time series are highly correlated. The main difference between the OLS model with backward elimination and that with LASSO is in the lags of the endogenous variable, the internal temperature, where the former has selected certain lags , while the later has selected all the lags. In terms of the exogenous variables, the OLS LASSO model has eliminated all of the outdoor temperature features as well as those of the CO_2 fan speed demand. It is possible that a different form of this exogenous feature would be more appropriate, one that would take into account the lagged values of past days and not only those of past hours, allowing the role of the thermal mass of the building to be captured in the models' structure.

Discussion

The aim of this work is to develop an efficient data-driven approach to model a building's thermal dynamics in response to various HVAC inputs.

The analysis indicates that the LASSO performs slightly worse in prediction but is able to reduce the number of parameters through the regularisation approach. This approach has its benefits as the model structure is defined at the same time as the model

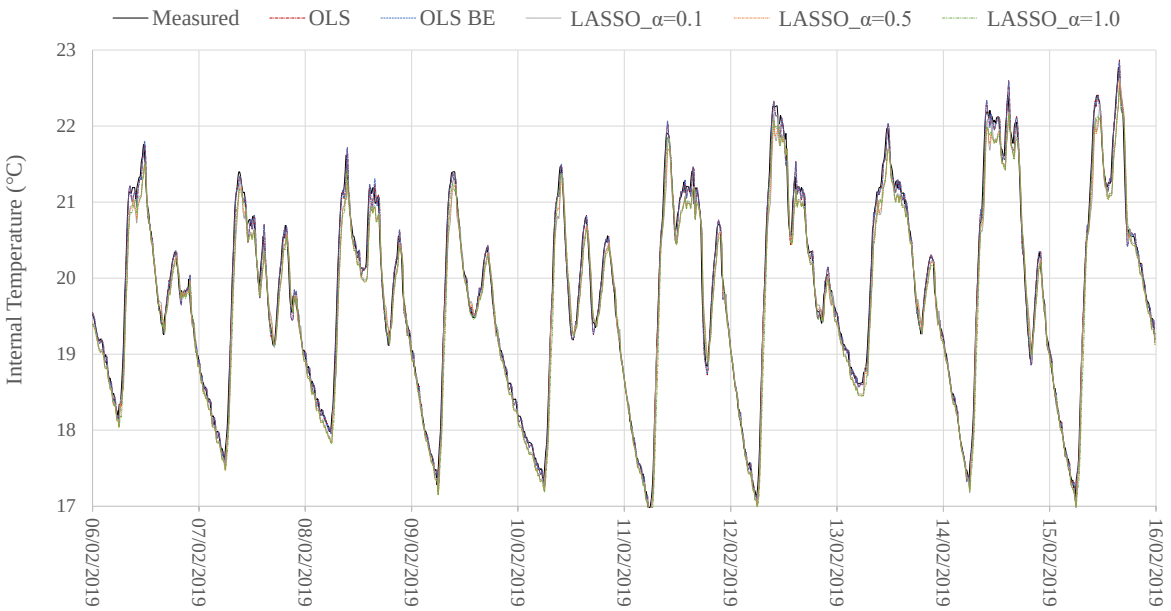

Figure 4: Model results for test data period

Table 1: Measures of error for model evaluation

Models	R^2	RMSE	CVRMSE (%)	MAE	MAPE (%)	MBE	NMBE (%)
OLS	0.99	0.1	0.60	0.1	0.45	0.0	0.09
OLS BE	0.99	0.1	0.60	0.1	0.44	0.0	0.09
OLS LASSO ($\alpha = 0.1$)	0.98	0.2	1.07	0.2	0.78	0.1	0.49
OLS LASSO ($\alpha = 0.5$)	0.98	0.2	1.09	0.2	0.83	0.1	0.61
OLS LASSO ($\alpha = 1.0$)	0.98	0.2	1.17	0.2	0.86	0.1	0.68

parameters.

In developing data-driven models of building system dynamics, one could encounter a variety of different system configurations. With the data-driven OLS LASSO approach, a model developer just needs to define the parameters of interest and the pertinent time scales of importance, i.e. how far back into the past to look, to define a model that could be used in control. This could be done offline to first determine which parameters are important to gather and then refit online with only a subset of parameters.

However, the results from this analysis indicate some issues with respect to the use of such models for control. The LASSO with an alpha value of 1.0 (LASSO α=1.0) found that the cooling valve, outside air temperature and CO_2 were not relevant to the building dynamics.

The outside air temperature not being a significant parameter is obviously not desirable as from building physics theory it is known that the outside air temperature will have an effect on the indoor temperature. The outdoor temperature not being relevant could be due to the few previous time lags considered, specifically 1 hour. This indicates that the selection of the bounds of the parameter space to be considered are very important.

With respect to the CO_2 Fan Speed Demand and the Cooling Valve Demand, given that the estimation was done in the winter period, it is possible that in this time period this parameters have little effect. Yet, it is possible, during other seasons, these parameter would have a large influence on the indoor temperature. This indicates the model would need to be re-estimated periodically.

Further the LASSO (α=1.0), found that only the heating valve demand at time $t - 4$ was relevant. This could potentially lead to poor control behaviour as the model would consider that the heating valve demand at other time lags would have no influence on the room temperature, leading to plausible gaps in the sequential predictions of the dependent variable, in this case the internal temperature of Teaching Room 1. The same would occur with the damper calculated demand.

The values of LASSO (α=0.1) also exhibit behaviour that would not be beneficial for control. The heat-

Table 2: Feature Coefficients

Features	OLS	OLS BE	OLS LASSO ($\alpha = 0.1$)	OLS LASSO ($\alpha = 0.5$)	OLS LASSO ($\alpha = 1.0$)
Cooling Valve Demand (t)	-0.0007	-0.0007	0	0	0
Cooling Valve Demand (t-1)	0.0000	0	0	0	0
Cooling Valve Demand (t-2)	-0.0002	0	0	0	0
Cooling Valve Demand (t-3)	0.0002	0	-0.0005	0	0
Cooling Valve Demand (t-4)	0.0007	0.0007	0	0	0
CO_2 Fan Speed Demand (t)	-0.0069	-0.0057	0	0	0
CO_2 Fan Speed Demand (t-1)	0.0023	0	0	0	0
CO_2 Fan Speed Demand (t-2)	0.0031	0.0028	0	0	0
CO_2 Fan Speed Demand (t-3)	-0.0034	0	0	0	0
CO_2 Fan Speed Demand (t-4)	0.0046	0.0025	0	0	0
Heating Valve Demand (t)	0.0036	0.0036	0.0122	0.0046	0
Heating Valve Demand (t-1)	-0.0037	-0.0035	-0.0033	0	0
Heating Valve Demand (t-2)	0.0000	0	-0.0053	0	0
Heating Valve Demand (t-3)	0.0018	0.0017	-0.0018	0	0
Heating Valve Demand (t-4)	0.0001	0	0.0042	0.0009	0.0046
Damper Calculated Demand (t)	0.0001	0	0.0024	0.0020	0
Damper Calculated Demand (t-1)	-0.0002	0	0.0005	0	0
Damper Calculated Demand (t-2)	0.0001	0	-0.0005	0	0.0010
Damper Calculated Demand (t-3)	0.0005	0.0006	-0.0024	0	0
Damper Calculated Demand (t-4)	0.0001	0	0.0023	0.0004	0.0010
Outside Air Temperature (t)	-0.0077	0	0	0	0
Outside Air Temperature (t-1)	0.0363	0.0142	0	0	0
Outside Air Temperature (t-2)	-0.0211	0	0	0	0
Outside Air Temperature (t-3)	0.0067	0	0	0	0
Outside Air Temperature (t-4)	-0.0107	-0.0110	0	0	0
Teaching Room 1 Temp (t)	1.3899	1.3917	0.9252	0.9262	0.9304
Teaching Room 1 Temp (t-1)	-0.0855	-0.0870	0.0020	0.0012	0.0006
Teaching Room 1 Temp (t-2)	-0.2768	-0.2668	0.0024	0.0012	0.0007
Teaching Room 1 Temp (t-3)	-0.0603	-0.0928	0.0026	0.0029	0.0011
Teaching Room 1 Temp (t-4)	-0.0241	0	0.0029	0.0032	0.0018
Teaching Room 1 Temp (t-5)	0.0067	0	0.0036	0.0037	0.0037
Teaching Room 1 Temp (t-6)	-0.0300	0	0.0040	0.0040	0.0040
Teaching Room 1 Temp (t-7)	0.0319	0	0.0043	0.0043	0.0043
Teaching Room 1 Temp (t-8)	0.0188	0.0329	0.0045	0.0045	0.0045
Teaching Room 1 Temp (t-9)	0.0106	0	0.0048	0.0047	0.0048
Teaching Room 1 Temp (t-10)	-0.0041	0	0.0050	0.0049	0.0049
Teaching Room 1 Temp (t-11)	0.0814	0.0833	0.0050	0.0050	0.0050
Teaching Room 1 Temp (t-12)	-0.0755	-0.1210	0.0050	0.0050	0.0049
Teaching Room 1 Temp (t-13)	-0.0588	0	0.0051	0.0050	0.0050
Teaching Room 1 Temp (t-14)	0.0507	0.0567	0.0052	0.0050	0.0050
Teaching Room 1 Temp (t-15)	0.0178	0	0.0050	0.0049	0.0049
Teaching Room 1 Temp (t-16)	0.0041	0	0.0046	0.0045	0.0045

ing value demands at times $t-1$, $t-2$, and $t-3$ are defined as negative values. This indicates that a heating value demand at these times reduces room temperature. This could also lead to undesirable control behaviour.

However, the OLS with backward selection exhibited much poorer performance in this respect, as there are many variables in the time history that set to zero and negative parameter values.

Therefore the LASSO approach would appear to be an efficient approach to developing a model structure as well as parameter identification, however further regularisation techniques may be needed to develop models that have good control performance.

The conclusions of this work are limited by the case study's small scope as it only considered a single room served by a single air handling unit in a single educational building. Nonetheless, the issues exposed for the parameter selection can generally be considered for other building energy systems. To provide more

general conclusions with respect to performance, this approach should be explored in a wide range of buildings across a variety of seasons.

Conclusions and Future Research

This work explored two feature selection techniques, the backwards elimination and LASSO regularisation, to estimate one-step ahead models of building thermal dynamics. The analysis found that across all metrics the LASSO approach performed worse than the backwards selection but the change in performance was very small. This indicates that the LASSO regularisation is an efficient approach to model selection.

The analysis did however indicate that the final model structures selected by both the backward selection and LASSO techniques could lead to poor control performance, due to negative values and lack of regularity in the time lags of exogenous variables, leading to plausible irregularities in the sequential form of the predictions.

Therefore, future work will focus on expanding the regularisation techniques to encourage selection of parameters with characteristics for stable control, extend the time history explored to include previous days, and examine the performance in a variety of buildings across various seasons.

Acknowledgement

This research was made possible by the support from the Engineering and Physical Sciences Research Council (EPSRC) for the FlexTECC: Flexible Timing of Energy onsumption in Communities Innovation Fellowship (grant EP/S001670/1).

References

Amiri, S. S., M. Mottahedi, and S. Asadi (2015). Using multiple regression analysis to develop energy consumption indicators for commercial buildings in the U.S. *Energy and Buildings 109*, 209–216.

Candanedo, L. M., V. Feldheim, and D. Deramaix (2017). Data driven prediction models of energy use of appliances in a low-energy house. *Energy and Buildings 140*, 81–97.

Chandrashekar, G. and F. Sahin (2014). A survey on feature selection methods. *Computers and Electrical Engineering 40*(1), 16–28.

Chen, Z., J. Freihaut, B. Lin, and C. Wang (2018). Inverse energy model development via high-dimensional data analysis and sub-metering priority in building data monitoring. *Energy and Buildings 172*, 116–124.

Draper, N. R. and H. Smith (2000). *Applied regression analysis* (3rd ed.). New York: John Wiley & Sons, Inc.

Fan, C., F. Xiao, and S. Wang (2014). Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy 127*, 1–10.

Geneidy, R. and B. Howard (2020). Contracted energy flexibility characteristics of communities: Analysis of a control strategy for demand response. *Applied Energy 263*(October 2019), 114600.

Guyon, I. and A. Elisseeff (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research 3*, 1157–1182.

Jović, A., K. Brkić, and N. Bogunović (2015). A review of feature selection methods with applications. *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings* (May), 1200–1205.

Liu, H. and M. Hiroshi (1998). *Feature selection for knowledge discovery and data mining*. Boston: Kluwer Academic.

Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control 34*(1), 1–12.

Ma, J. and J. C. Cheng (2016). Identifying the influential features on the regional energy use intensity of residential buildings based on Random Forests. *Applied Energy 183*, 193–201.

Montgomery, D. C. (2011). *Applied statistics and probability for engineers* (5th ed.). hoboken New Jersey: John Wiley.

Rätz, M., A. P. Javadi, M. Baranski, K. Finkbeiner, and D. Müller (2019). Automated data-driven modeling of building energy systems via machine learning algorithms. *Energy and Buildings 202*, 109384.

Suryanarayana, G., J. Lago, D. Geysen, P. Aleksiejuk, and C. Johansson (2018). Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. *Energy 157*, 141–149.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

Vu, D. H., K. M. Muttaqi, and A. P. Agalgaonkar (2015). A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Applied Energy 140*, 385–394.

Wang, W., T. Hong, N. Xu, X. Xu, J. Chen, and X. Shan (2019). Cross-source sensing data fusion for building occupancy prediction with adaptive lasso feature filtering. *Building and Environment 162*(June), 106280.

Zadeh, L. A. (1956). On the identification problem. *IRE Transactions on Circuit Theory CT-3*, 277–281.

Zhang, L. and J. Wen (2019). A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy and Buildings 183*, 428–442.