# Using regression analysis to predict the future energy consumption of a supermarket in the UK

CrossMark

M.R. Braun [a,*], H. Altan [b], S.B.M. Beck [a]

[a] Department of Mechanical Engineering, The University of Sheffield, Sheffield S1 3JD, UK
[b] Faculty of Engineering & IT, The British University in Dubai, Dubai, United Arab Emirates

## HIGHLIGHTS

- Energy consumption of supermarket depends more on temperature than humidity.
- Multiple regression analysis is a flexible tool to consider for energy use prediction.
- Results show dramatic reduction in gas use and modest increase in electricity use.

## ARTICLE INFO

## ABSTRACT

The change in climate has led to an interest in how this will affect the energy consumption in buildings. Most of the work in the literature relates to offices and homes. However, this paper investigates a supermarket in northern England by means of a multiple regression analysis based on gas and electricity data for 2012.

The equations obtained in this analysis use the humidity ratio derived from the dry-bulb temperature and the relative humidity in conjunction with the actual dry-bulb temperature. These equations are used to estimate the consumption for the base year period (1961–1990) and for the predicted climate period 2030–2059.

The findings indicate that electricity use will increase by 2.1% whereas gas consumption will drop by about 13% for the central future estimate. The research further suggests that the year 2012 is comparable in temperature to the future climate, but the relative humidity is lower. Further research should include adaptation/mitigation measures and an evaluation of their usefulness.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

## 1. Introduction

With the founding of the IPCC in 1988 [1] the idea of anthropogenic climate change really entered the scientific arena so that, currently, the vast majority of researchers working in this area believe that the climate is changing and that this is fundamentally man-made [2]. That the issue of global warming has also reached politics is evident by the coming into force of the UN framework convention on climate change [3] in 1994 [4]. All of this has added to the interest in assessing the impact of climate change on various aspects of society, including on energy consumption in buildings.

Some of those assessments examine specific countries such as the UK. Jenkins et al. [5], for example, use a software model of a four-story office building to investigate five locations in the UK to see how the change in climate will affect the energy demand for heating and cooling in 2030. These researchers find that the energy demand, although in part location dependent, is primarily heating dominated. Their study also includes the assumption that office equipment and lighting will be more efficient (so producing less waste heat) in the future which will increase the demand for heating. However, they conclude that the temperature increase due to climate change will mitigate this to a degree. Gupta and Gregg [6] evaluate the effect of climate change on four types of dwelling located in Oxford, UK, by means of the simulation software IES. They find that thermal discomfort will rise significantly with climate change, especially in flats.

A number of other studies are summarized by Li et al. [7] who point out the two main approaches: the degree-day method and simulation techniques. Most of the papers in that review study office buildings and homes. The authors find that the predicted warming will result in a reduced heating load and an increased cooling load. This translates into a reduction in energy use for

---

colder climates and an increase in electricity consumption for warmer climates.

In addition to the degree-day method and simulation, other approaches have been used. One example is the paper by Schrock and Claridge [8] in which the authors use a simple regression model of the ambient temperature to investigate a supermarket's electricity use. The use of multiple linear regression analysis allows the inclusion of any desired variable. This technique is used by Lam et al. [9] who study office buildings in different climates in China. These researchers include 12 input variables covering building parameters, building loads and the HVAC system in their regression model and find that predictions largely agree with building software simulation. Another example is Chung et al. [10] who investigate the energy use intensity of supermarkets by means of such diverse variables as operational schedule, number of customers, lighting control, employee behaviour and maintenance factors, but explicitly exclude outdoor climate. Braun et al. [11] employ multiple regression analysis to investigate timer settings, night cover effectiveness together with indoor and outdoor temperature and humidity on the electricity consumption of a supermarket. The more complex principal component analysis is used by Lam et al. in [12] for office buildings. This technique allows the same flexibility as multiple regression analysis, but is not restricted by its underlying assumptions (for more details see 2.3 Multiple linear regression analysis).

Owing to their refrigerated shelves, supermarkets are quite different from other commercial buildings. Consequently, there are a number of documents published relating to modelling their energy consumption. One example is Suzuki et al. [13] who model the refrigeration system and energy flow of heat sources of a supermarket in Japan for one-hour increments. The authors find that the refrigeration equipment accounts for about 60% of the total energy demand and that the air leakage of the open refrigerated shelves has a considerable effect on this demand. The paper by Arias and Lundqvist [14] presents their software *CyperMart* which focuses on different refrigeration systems. Although the software uses also climate data as inputs, the intended users are not those researching climate change impact, but designers and technicians. The work of Bahman et al. [15] uses a moisture balance equation and the humidity ratio $w$ to infer the indoor relative humidity in a supermarket. This, in turn, is used to simulate the energy use. Their results suggest that the indoor relative humidity is strongly correlated with the total energy consumption, i.e. the lower the indoor humidity, the lower the total energy use.

There are two documents which relate to the carbon footprint of supermarkets in the UK. The first is a report by ENDS Carbons [16], which looks at supermarkets as a whole, from direct to indirect $CO_2$ emissions. However, it does not quantify the impact of climate change on supermarkets. The second document is a paper by Jenkins [17] in which the author uses a software model to evaluate different carbon-saving measures of a "standard UK supermarket". This research does not explicitly model the refrigeration systems which may have the effect of insufficiently capturing the main difference between supermarkets and other retail buildings.

The review above suggests that there will be an impact of the changing climate on the energy consumption of supermarkets and that relative humidity is likely a parameter which needs including in addition to outside temperature (see for instance [15,18]). Although some researchers examined supermarkets' $CO_2$ emissions, none have reported on the impact of future climate on their energy use. Therefore this paper focuses on predicting the change in energy consumption of a grocery supermarket for the 2040s by using the outside temperature and relative humidity data to perform a multiple regression analysis. The 2040s (rather than the end of the century) were chosen because (a) the lifespan of a refrigeration system is typically 15–20 years and (b) this research is aimed to be relevant to present day strategic decision makers in designing and operating current and prospective supermarket buildings and systems, and whose own lifetime might very well include the 2040s.

## 2. Study method

The supermarket studied and the methodology of the analysis and modelling is detailed in this section. As Fig. 1 indicates, this assessment is based on the actual consumption data, dry-bulb temperature and relative humidity records for 2012. This data was divided into two data sets to be used in a multiple linear regression analysis to generate two equations, one for electricity and one for gas. Thereafter these equations were used to estimate the consumption for the base period (1961–1990) which then was compared with the estimated consumption for the future period (2030–2059, also called the '2040s').

The study method used here may have certain limitations because it does not consider any other weather parameters, such as solar radiation or wind, or any future technical advances and building improvements. Furthermore it assumes that any change in footfall is negligible. Notwithstanding that, this method yields meaningful results in an easily realisable way.

### 2.1. Supermarket

The supermarket (see Fig. 2), which opened in July 2010, is located in the UK Yorkshire and Humber region, close to the city of Hull. It is at the larger end of the mid-range store size with a sales area of 1266 $m^2$ (see also Table 1) and an electric energy use density for 2012 of about 460 kW h/$m^2$ pa which, according to Tassou et al. [19], is about half the expected value. The supermarket sells mainly food and some general merchandise. In addition, it also has a café/restaurant and a small bakery.

The main energy consumers are the 240 kW condensing gas boiler (which serves the cold aisle heating, the general heating and the hot water system) and the two remote refrigeration R404/$CO_2$ plants of nominally 80 kW and 60 kW cooling capacity. These two plants are responsible for refrigerating about one third of all the shelves on the sales floor and the cold room. The three freezer cabinets in the sales area are self-contained and the freezer room is connected to two small freezer units outside.

The breakdown of electricity use in Fig. 3 is based on the consumption data from sub-meters for the first half of 2012 and is typical for supermarkets [20,21]. This figure shows that about half of the electricity consumption is made up of essentially weather independent loads such as lights (lights are also fitted to the refrigerated shelves). The consumption of the other half, i.e. HVAC and refrigeration packs, is more directly related to the weather.

### 2.2. Weather and consumption data

The datasets for the regression analysis are based on consumption and weather data for the whole year 2012. The temperature and relative humidity was downloaded in 15-min increments from a sensor situated on the north side of the supermarket. The gas and electricity consumption for the whole supermarket was readily available from the company's energy website. This data was downloaded in hourly readings and then summed for each week.

The temperature and humidity data for the base year was downloaded from the MET office website [22]. These values are monthly, long term averages for the period of 1961 to 1990 for the 25 km square containing Hull. The same two weather variables were also obtained for the same square from the UKCP09 website [23] by downloading monthly predictions for the high emissions
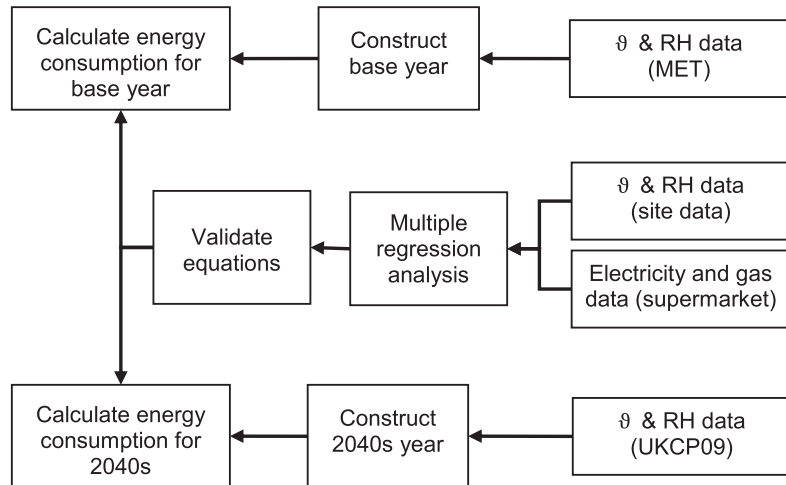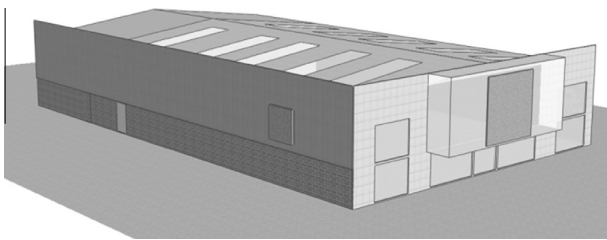
**Fig. 1.** Method flowchart of main steps.



**Fig. 2.** Supermarket for case study.

**Table 1**
Supermarket dimensions.

| Area name | Width (m) | Length (m) | Area surface (m²) |
|---|---|---|---|
| Total area of building | 32 | 57 | 1824 |
| Sales area (incl. café/restaurant) | 32 | 40 | 1266 |
| Café/restaurant | 7 | 16.5 | 115.5 |

Remarks: sales area has a small recess.
Estimated height: 8 m at sides, 10 m in middle.

scenario for the 2040s (i.e. the period from 2030 to 2059). After that the 10th, 50th and 90th percentile were extracted so that not only the central estimate (50th percentile) but also lower and upper limits could be calculated. These monthly values were entered as the central week in their respective month of 2012. The values in between were generated by the Matlab [24] interp1 function using the spline option, obtaining the smooth graphs shown in Fig. 4.

Actual weather data can be split into two components: deterministic (periodic) and stochastic (random) [25]. From the graphs in Fig. 4, it can be seen that only the actual weather traces include a random component, but not the weather for the base and predicted future climate years. This is particularly apparent in the left-hand panel for temperature, in which the actual temperature trace follows the central estimate trace quite well – being sometimes below and sometimes above it. The notable exception is when the actual temperature dips and rises quite sharply from late January to mid February. In the right-hand graph the actual relative humidity does not follow any of the other curves, but stays below them with only a modest amount of variation due to seasonality. Fig. 4 is also interesting from the data coverage point of view because it shows that the required temperature range for forecasting is being covered by the actual temperature, but not by the relative humidity. The ramifications of this will be discussed later.

As relative humidity is also a function of the dry-bulb temperature, it was transformed into a value which is directly meaningful without reference to the dry-bulb temperature. According to Lawrence [26] relative humidity is commonly thought of as either the ratio of actual water vapour pressure to the saturation vapour pressure or as the ratio of the air humidity ratio $w$ to the saturation
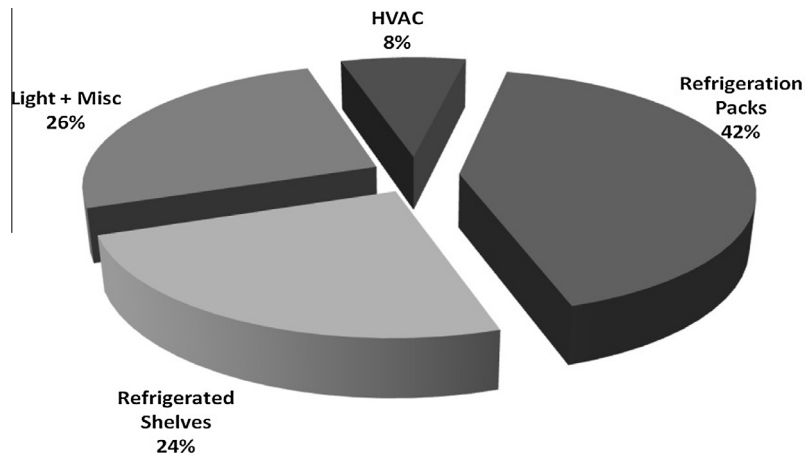


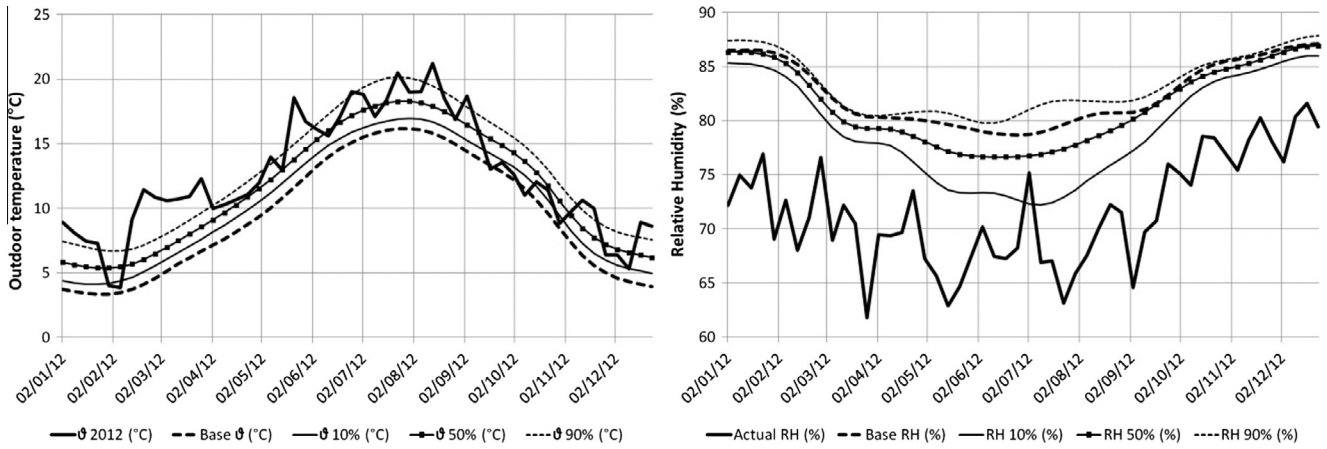**Fig. 3.** Breakdown of electricity consumption.

**Fig. 4.** Temperature and relative humidity plots.

mixing ratio $w_s$ at a given temperature and pressure (also called the 'degree of saturation' [27]). He also points out that these two definitions are not interchangeable. Here, the second definition is assumed to be correct (for another example see [15], Fig. 5, where the authors calculate relative humidity from the humidity ratio and ambient temperature). An equation for the saturation mixing ratio $w_s$ was derived from Table 2 on page 6.3 in [27] by fitting an exponential function to the values for $w_s$ for the temperatures between 0 °C and 35 °C.

### 2.3. Multiple linear regression analysis

Multiple linear regression analysis seeks to establish a relationship between a dependent variable (in this case the energy consumption) and two or more independent variables (the predictors) in the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon$$

In this equation $\beta_0 \ldots \beta_n$ are the regression coefficients to be estimated based on a record of observations. This is normally done by curve fitting based on the least square method with the aim of

**Table 2**
Correlation coefficients for electricity dataset.

|  | Electricity | $\vartheta$ | $\vartheta^2$ | $w$ | $w^2$ |
|---|---|---|---|---|---|
| Electricity | 1 | | | | |
| $\vartheta$ | 0.784 | 1 | | | |
| $\vartheta^2$ | 0.860 | 0.982 | 1 | | |
| $w$ | 0.842 | 0.968 | 0.971 | 1 | |
| $w^2$ | 0.884 | 0.947 | 0.971 | 0.992 | 1 |
| $\vartheta \cdot w$ | 0.875 | 0.975 | 0.994 | 0.989 | 0.991 |

minimizing the difference between the observed and estimated values. The predictors should have little or no correlation with each other (i.e. the correlation coefficient should be less than 0.7 [28]) to avoid problems caused by multicollinearity. The last term in the equation, $\varepsilon$, is referred to as the residual (or fitted error) and is used for testing the overall significance (*F*-test) of the equation and the significance of each regression coefficient (*t*-test). In order to obtain valid results from these tests the residual $\varepsilon$ has to be normally and independently distributed, with a mean of zero and a constant variance of $\sigma^2$ [29]. This is verified by what is called a
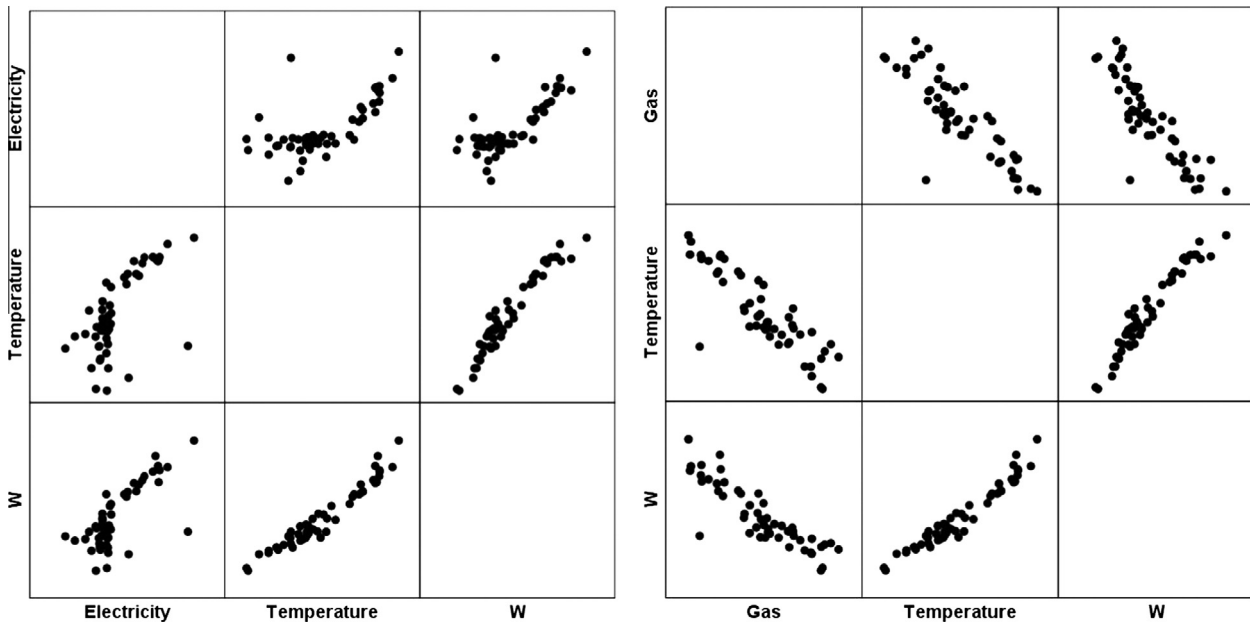


**Fig. 5.** Scatter plot matrices for original gas and electricity data.

residual analysis. This analysis may also lead to the elimination of data outliers.

Another important indicator is the coefficient of determination, $R^2$, which not only indicates the goodness of fit, but can also be interpreted as the amount of variation of the dependent variable explained by the regression equation [28].

Before a regression model is selected, it is advisable to look at a scatter plot matrix of the dependent variable and predictors to see if linear regression is appropriate and, if so, what model should be adopted. Fig. 5 shows such matrices for the original data sets. The left-hand panel, which displays the scatter plots for electricity, temperature and humidity ratio, indicates that electricity use is quite stable for the first two thirds before it starts to rise. Here the incorporation of higher order terms seems advisable. The right-hand panel is for the gas consumption dataset. It shows a more linear relationship between the independent variables and the gas consumption. Here a linear equation might be sufficient, but to take care of any curvature a second order term was incorporated. Both scatter plot matrices also indicate a strong correlation between the independent variables and the possibility of outliers.

Based on the reasoning above, the following second-order model with interaction was adopted:

$$y_{electricity\ or\ gas} = \beta_0 + \beta_1\vartheta + \beta_2\vartheta^2 + \beta_2 w + \beta_4 w^2 + \beta_5\vartheta w + \varepsilon$$

Another preparatory step is to test the predictors for multicollinearity by, for instance, calculating their correlation coefficients $R$. This has been done and is summarized in Tables 2 and 3. Both tables show a very high level of correlation between all predictors so that it is expected that only one or two predictors remain after the multiple regression analysis.

The actual regression analysis was performed iteratively with the software package IBM SPSS Statistics, version 21 [30]. The option chosen for both datasets was stepwise (inclusion $p$-value = 0.50 and exclusion $p$-value = 0.55). During these iterations, outliers of the residual terms were eliminated, because the statistics calculated during the linear regression analysis rely on a normal distribution of these error terms. The extent to which these terms were eliminated is indicated in Fig. 6 of the standardized error terms. There the left-hand side displays the distribution with the outliers and the right-hand panel without them. The outliers included data points for holidays, such as the Christmas period when the store has irregular opening hours and staff work overnight. When examining the residuals for the gas data set, autocorrelation of the error for the first half of 2012 was noticed. As this also indicates a violation of the underlying assumption of linear regression [29] it was reduced as much as possible, by deleting a limited number of data points at the beginning of the year. After treating the data this way, 44 data points in the electricity data set and 39 in the gas consumption data set remained.

These treated data sets yielded the following equations:

$$\hat{y}_{electricty} = 13030.1 - 841262.5w + 54348895.8w^2 + 13157.1\vartheta \cdot w$$
$$_{t\text{-statistics:}} \qquad (33.7) \qquad (-7.8) \qquad (6.8) \qquad (3.7)$$

Statistics : $R = 0.976, R^2_{adj} = 0.949, F = 268,$

Durbin − Watson = 1.733

**Table 3**
Correlation coefficients for gas dataset.

|  | Gas | $\vartheta$ | $\vartheta^2$ | $w$ | $w^2$ |
|---|---|---|---|---|---|
| Gas | 1 | | | | |
| $\vartheta$ | −0.919 | 1 | | | |
| $\vartheta^2$ | −0.924 | 0.988 | 1 | | |
| $w$ | −0.929 | 0.959 | 0.971 | 1 | |
| $w^2$ | −0.910 | 0.936 | 0.962 | 0.994 | 1 |
| $\vartheta \cdot w$ | −0.927 | 0.975 | 0.992 | 0.991 | 0.988 |

and

$$\hat{y}_{gas} = 17315.2 - 1424062.7w$$
$$_{t\text{-statistics:}} \qquad (25.7) \qquad (-15.4)$$

Statistics : $R = 0.930, R^2_{adj} = 0.864, F = 236,$

Durbin − Watson = 1.179

The residual analysis showed that:

– For electricity the error was normally distributed and had constant variance. The Durbin–Watson statistic suggested that the errors were not autocorrelated.
– For gas the error was normally distributed and had constant variance. The Durbin–Watson statistic suggested that the error terms may have been autocorrelated.

As Tables 2 and 3 indicate, there are close correlations between the predictors, which means that just using temperature could also yield acceptable equations. For instance, a regression equation using the first and second order temperature terms for electricity consumption gives an $R = 0.958$ ($R^2_{adj} = 0.918$). The $t$-statistics show that both predictors are significant. However, the Durbin–Watson statistic (=1.167) is not as good as for the equation given above. A gas equation based on temperature alone performs best with just a square term ($R = 0.926$, $R^2_{adj} = 0.854$). Here the $t$-statistics suggest that the linear term is not significant. Also in this case the Durbin–Watson statistic (=0.857) was not as good as for the equation above. This drop in $R$ and the increased autocorrelation leads to the conclusion that the equation including humidity ratio terms perform better, especially if one considers that [29], page 476, states that autocorrelation indicates the presence of an omitted predictor variable.

### 2.4. Model evaluation

Fig. 7 compares the predicted values of the regression analysis with the actual values from the datasets and shows that both equations perform satisfactorily. This figure can also be interpreted as the visualisation of the value $R$ given above.

The attempt to access data for 2013 to compare the actual consumption with the predicted energy use proved only partially successful. It was found that the data for the onside temperature and humidity sensor was only available from the middle of October 2013 onwards. In addition to this, technical problems with the gas boiler meant that it was replaced with a more modern, higher capacity boiler which made it impossible to use the available data for validating the gas consumption equation. The comparison of the actual and predicted electricity consumption proved more successful and is summarized in Table 4. As other researchers have done before (e.g. [31]), here also the two error statistics NMBE (Normalised Mean Biased Error) and CVRMSE (Coefficient of Variation of the Root-Mean-Square Error) have been calculated based on Table 4. The NMBE is −2.6% and indicates that the estimated values are, on average, slightly too low. The CVRMSE is just below 3.8%, suggesting that the estimation equation presented above works reasonably well over the considered period.

The error estimate based on error propagation is quite difficult to do because some of the information is missing. For instance, no information regarding the error of the actual consumption data and the weather sensor was available. For the base year period, the Met Office gave an RMSE of 0.94 °C [32]. An error for the relative humidity was not given. To estimate the error introduced by the regression prediction the formula in [29], page 99 was implemented in Matlab and the error for the yearly electricity consumption was found to be less than 3% and for gas use 4.6% or less (at a 95% confidence level).
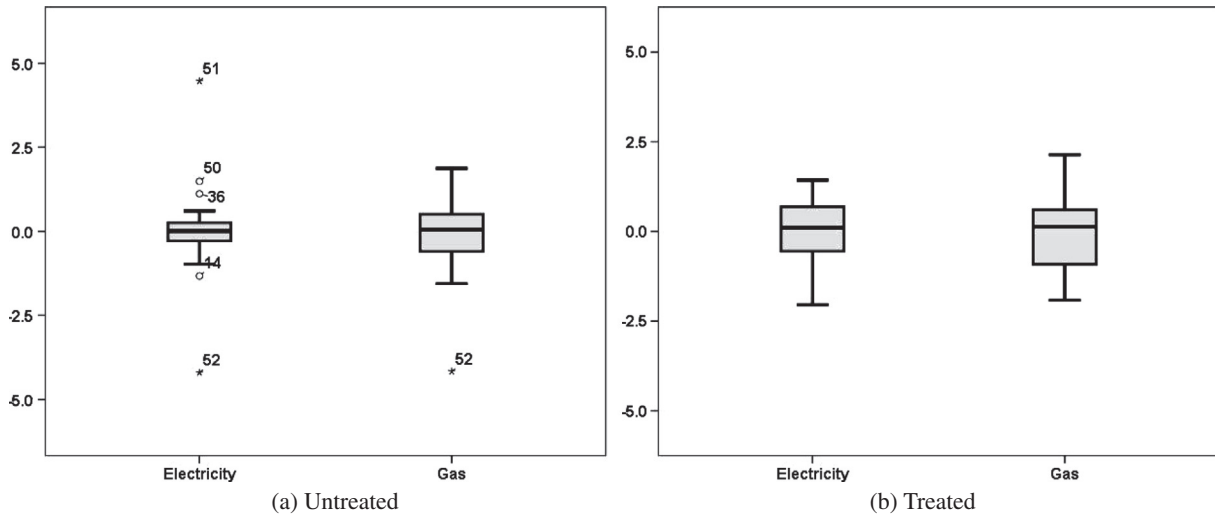
(a) Untreated                                          (b) Treated

**Fig. 6.** Box plot for standardized error terms.



Electricity                                               Gas
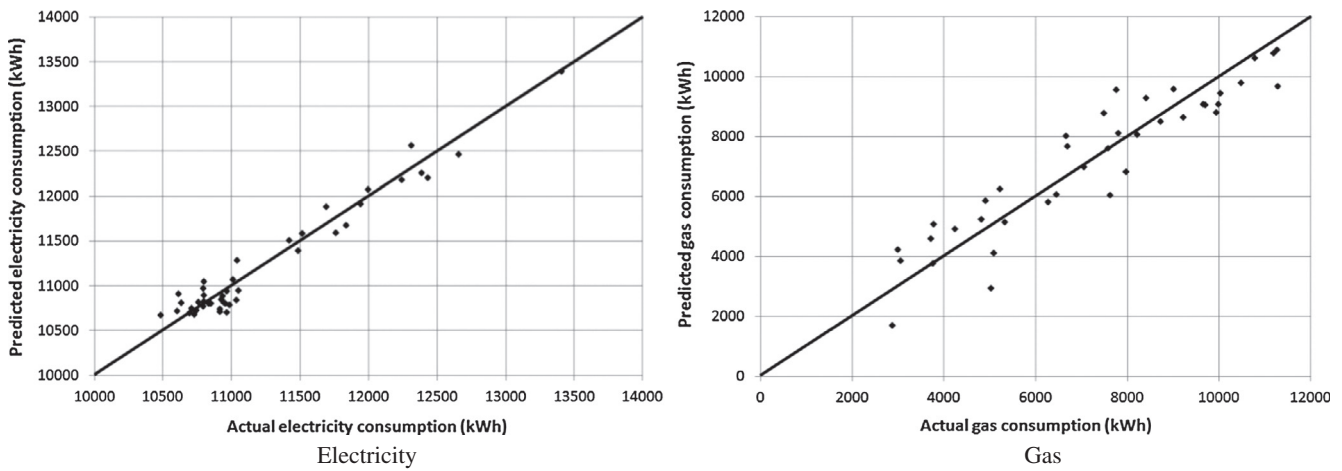
**Fig. 7.** Comparing prediction with actual values.

**Table 4**
Comparing actual electricity consumption with predicted consumption.

| W/C | Actual consumption (kW h) | Estimated consumption (kW h) | Difference (kW h) |
| --- | --- | --- | --- |
| 14/10/2013 | 11,180 | 11129.1 | −50.95 |
| 21/10/2013 | 11,443 | 11387.8 | −55.19 |
| 28/10/2013 | 10,782 | 10812.1 | 30.11 |
| 04/11/2013 | 10,686 | 10672.6 | −13.44 |
| 11/11/2013 | 10,938 | 10755.7 | −182.35 |
| 18/11/2013 | 10,669 | 10698.9 | 29.89 |
| 25/11/2013 | 10,968 | 10709.2 | −258.79 |
| 02/12/2013 | 11,042 | 10754.3 | −287.74 |
| 09/12/2013 | 11,662 | 10712.5 | −949.48 |
| 16/12/2013 | Christmas period (disregarded) | | |
| 23/12/2013 | | | |
| 30/12/2013 | | | |
| 06/01/2014 | 11,301 | 10694.1 | −606.94 |
| 13/01/2014 | 11,130 | 10652.1 | −477.93 |
| 20/01/2014 | 11,303 | 10644.4 | −658.59 |

To find out whether the change in relative humidity or in temperature was more important, a sensitivity analysis was performed which is summarized in Table 5. This table indicates that the temperature has a far greater influence on both gas and electricity consumption than humidity.
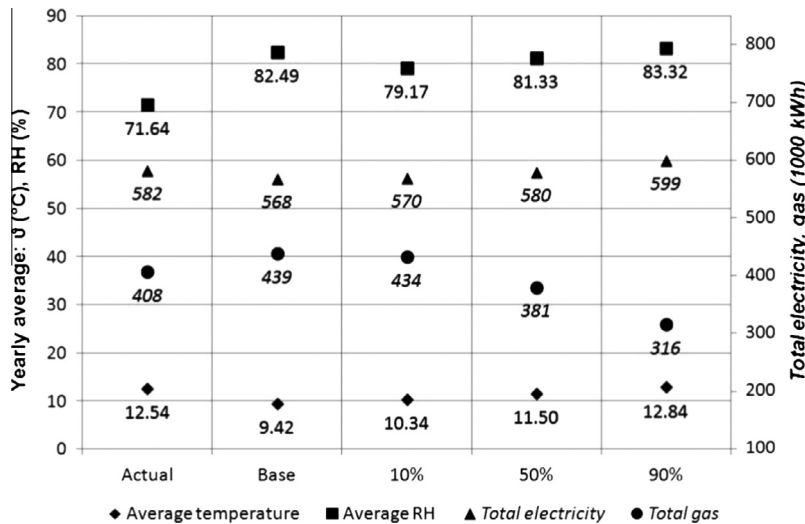
## 3. Results

After establishing that both equations perform reasonably well, they were used to estimate the consumption for the base year and for the 2040s (central estimate, the 10th and 90th percentile). The

**Table 5**
Sensitivity analysis.

| | | $\vartheta_{10\%}$ | $\vartheta_{50\%}$ | $\vartheta_{90\%}$ | \|Difference\| |
|---|---|---|---|---|---|
| Electricity (kW h) | $RH_{10\%}$ | 570,284 | | 595,235 | 24,951 |
| Gas (kW h) | | 433,528 | | 328,719 | 104,809 |
| Total (kW h) | | 1,003,812 | | 923,954 | 79,858 |
| Electricity (kW h) | $RH_{50\%}$ | | 580,431 | | |
| Gas (kW h) | | | 380,922 | | |
| Total (kW h) | | | 961,353 | | |
| Electricity (kW h) | $RH_{90\%}$ | 576,038 | | 599,339 | 23,301 |
| Gas (kW h) | | 389,404 | | 316,328 | 73,076 |
| Total (kW h) | | 965,442 | | 915,667 | 49,775 |
| Electricity (kW h) | \|Difference\| | 5754 | | 4104 | |
| Gas (kW h) | | 44,124 | | 12,391 | |
| Total (kW h) | | 38,370 | | 8287 | |
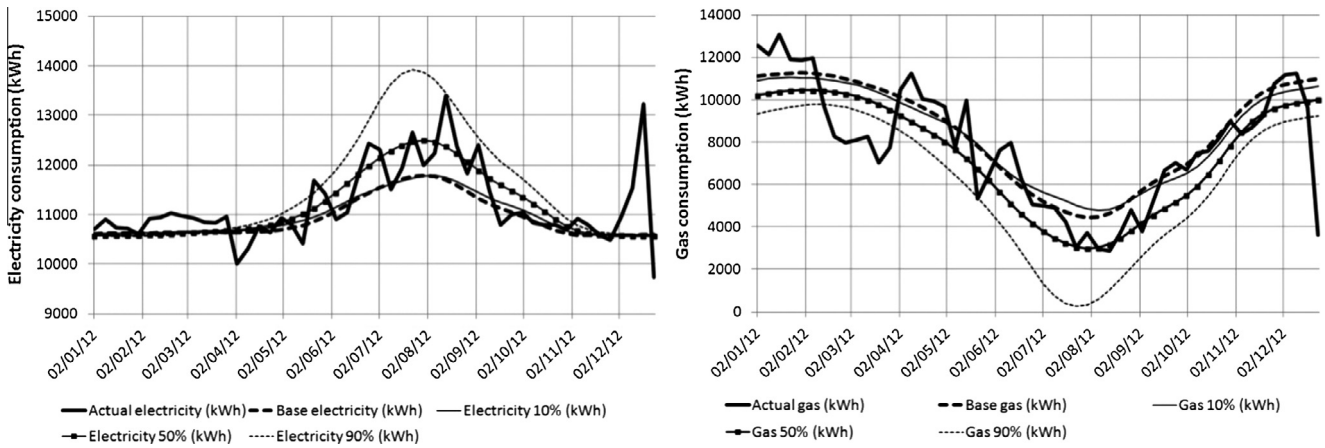


**Fig. 8.** Electricity and gas consumption.



**Fig. 9.** Summary plot.

results are plotted in Fig. 8. In both panels the base year and the graph for the 10th percentile are very close together. This can be explained by examining Fig. 4 which shows that the base year temperature is below the 10th percentile temperature, but that the relative humidity for the base year is considerably higher. The estimated consumption graphs in the left-hand plot are very close together during the winter and early spring months and fan out afterwards when the temperature increases. This may be due to

the square term in the electricity equation. The right-hand plot shows a relatively constant distance between all estimate graphs.

Fig. 9 summarizes all the results on a yearly basis. It shows that the temperature in 2012 is already at the level expected for the 2040s, but that the relative humidity is about 12% below the central estimate. It also shows that the average temperature in 2012 was about 3 °C above the long term average from 1961 to 1990. Comparing the electricity consumption of the base year with the

2040s, one finds that the central estimate is a 2.1% increase with only a slight increase (0.4%) at the lower limit and a rise of 5.5% as an upper limit. For the gas use, one finds a decrease ranging from about 1% to about 28% with about 13% being the central estimate.

## 4. Discussion and conclusions

This work describes how the energy consumption of a supermarket in northern England is expected to change for the period from 2030 to 2059 (the '2040s'). The basis for this assessment was the gas and electricity consumption as well as the outside temperature and relative humidity data for 2012. The regression equation derived from this data was used to estimate the consumption based on the long-term averages from 1961 to 1990, which then was compared with predictions for the 2040s. Based on this, the electricity consumption is thought to rise by up to 5.5% with 2.1% being the central estimate. The gas consumption is estimated to fall by up to 28% (13% central estimate). This is in line with results reported in [7] where the reviewers found studies for central and northern Europe suggesting that a decrease in heating would dominate there.

The results are based on two equations from a multiple regression analysis which make use of the air humidity ratio $w$. This approach is similar to the one used in [15] because in that study this ratio is also derived from temperature and relative humidity data and then used to estimate the energy consumption of a supermarket. The regression equation derived above can explain about 95% of the variation of the electricity demand and 86% of the gas use. If only temperature terms were allowed, this would drop, albeit only slightly, to 92% and 85%. This and the increase in autocorrelation can be used to argue that the equations applied in this work have a higher explanative power and are to be preferred to temperature terms only.

The equation for electricity incorporates $w$, $w^2$ and $w \cdot \vartheta$ terms. The coefficient for $w$ is negative whereas the coefficients for the other two terms are positive. When examining their individual contribution to the overall response, it can be found that the influence of the linear term diminishes with increasing $w$. Thus the virtually flat part of the predicted electricity consumption at the beginning, gives way to a steeper rise for higher values of the humidity ratio $w$. This mimics the dependency of the electricity use on $w$ very effectively, as seen in the left-hand panel in Fig. 5. The same scatter plot matrix also shows a close relationship between $w$ and $\vartheta$ so that it may be said that the linear term in the electricity equation is related to the heating effort which decreases as temperature increases. The other terms may capture the relationship between the cooling load and the increase in temperature. The regression equation for gas uses only a linear term of $w$ with a negative coefficient. This is also consistent with the right-hand panel of Fig. 5 and captures the decreasing heating demand with rising temperature.

Comparing the climate variables, one finds that the yearly temperature average of the actual measurements is over 30% (or 3.12 °C) higher than the base year. Notwithstanding that, almost the whole temperature range of future climate values is covered. On the other hand, the yearly relative humidity average (negatively) deviates by only about 13% from the base year, but the actual relative humidity does not cover the range of the base and future humidity estimates. Consequently the consumption estimates are based on extrapolating their relationship with relative humidity (and therefore may introduce unknown errors). The actual temperature data indicates that the year 2012 was warmer than average, which would be consistent with a lower than average relative humidity because warmer air can hold more moisture. Generally speaking, the UKCP09 estimates predict a climate which is warmer than the base year but with a similar relative humidity.

The work presented here will be expanded by, for instance, researching other supermarkets throughout Britain. The selection of locations could be modelled on the approach in [5] where the researchers studied one location in Scotland, three in England and one in Wales. Such a study could also explore whether including other predictors, such as other weather parameters and footfall, reduces autocorrelation. It might also be worthwhile to explore different modelling techniques using, say, principal component analysis (as in [12]) which offers the advantage of not requiring normally distributed error terms. Another approach could be the software modelling mentioned in [7] because this could also be used to examine energy saving options which would be of interest to supermarket companies. Finally, the impact of the recently published IPCC WG 1 report [33], on this work should be assessed in future research.

## References

[1] IPCC. History. n.d. <http://www.ipcc.ch/organization/organization_history.shtml#.UgtdAtLVCmg> [accessed: 14.08.13].
[2] Cook J, Nuccitelli D, Green SA, Richardson M, Winkler B, Painting R, et al. Quantifying the consensus on anthropogenic global warming in the scientific literature. Environ Res Lett 2013;8.
[3] United Nations. United Nations framework convention on climate change. New York: United Nations; 1992.
[4] United Nations Framework Convention on Climate Change. Introduction to the Convention 2013. <http://unfccc.int/essential_background/convention/items/6036.php> [accessed: 14.08.13].
[5] Jenkins D, Liu Y, Peacock AD. Climatic and internal factors affecting future UK office heating and cooling energy consumptions. Energy Build 2008;40(5):874–81.
[6] Gupta R, Gregg M. Using UK climate change projections to adapt existing English homes for a warming climate. Build Environ 2012;55:20–42.
[7] Li DHW, Yang L, Lam JC. Impact of climate change on energy use in the built environment in different climate zones – a review. Energy 2012;42(1):103–12.
[8] Schrock DW, Claridge DE. Predicting Energy Usage in a Supermarket, in: 6th Symposium on Improving Building Systems in Hot and Humid Climates. Dalas; 1989.
[9] Lam JC, Wan KKW, Liu D, Tsang CL. Multiple regression models for energy use in air-conditioned office buildings in different climates. Energy Convers Manage 2010;51(12):2692–7.
[10] Chung W, Hui YV, Lam YM. Benchmarking the energy efficiency of commercial buildings. Appl Energy 2006;83(1):1–14.
[11] Braun MR, Altan H, Beck SEB. Using IBM SPSS statistics to identify predictors of electricity consumption in a UK supermarket. In: Energy and sustainability 2013. Bucharest: The Wessex Institute of Technology; 2013.
[12] Lam JC, Wan KKW, Cheung KL, Yang L. Principal component analysis of electricity use in office buildings. Energy Build 2008;40(5):828–36.
[13] Suzuki Y, Yamaguchi Y, Shiraishi K, Narumi D, Shimoda Y. Analysis and modeling of energy demand of retail stores. In: Sydney 2011: 12th Conference of International Building Performance Simulation Association.

[14] Arias J, Lundqvist P. Modelling and experimental validation of advanced refrigeration systems in supermarkets. Proc Ins Mech Eng E, J Process Mech Eng 2005;219(E2):149–56.

[15] Bahman A, Rosario L, Rahman MM. Analysis of energy savings in a supermarket refrigeration/HVAC system. Appl Energy 2012;98:11–21.

[16] ENDS Carbon, UK Supermarkets, 2009 Carbon Benchmark Report, ENDS Carbon: London; 2009.

[17] Jenkins DP. Using dynamic simulation to quantify the effect of carbon-saving measures for a UK supermarket. J Build Perform Simulat 2008;1(4):275–88.

[18] Howell R, Rosario L, Bula A. Effects of indoor relative humidity on refrigerated display cases performance. In: Proceedings of CLIMA 2000 Conference. Bruselas. Bélgica; 1997.

[19] Tassou SA, Ge Y, Hadawey A, Marriott D. Energy consumption and conservation in food retailing. Appl Therm Eng 2011;31(2–3):147–56.

[20] Schraps S. Energiekostenreduzierung in Supermärkten. Kälte Klima Aktuell 2005;2:26–30.

[21] Arias J. Energy Usage in Supermarkets: Modelling and Field Measurements. 2005. Kungliga Tekniska Hogskolan (Sweden) DrTechn., Kungliga Tekniska Hogskolan (Sweden).

[22] Met Office. UKCP09: Download data sets. n. d. <http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/download/index.html>. [accessed: 09.09.13].

[23] Met Office, UKCIP, British Atmospheric Data Centre, Newcastle University, University of East Anglia, Environment Agency, et al. UK Climate Projections. 2013, <http://ukclimateprojections-ui.defra.gov.uk/ui/admin/login.php>. [accessed: 21.08.13]

[24] MathWorks, MATLAB 7, Getting Started Guide. 2011, The MathWorks: Natick, USA.

[25] Yoshida H, Terai T. An ARMA type weather model for air-conditioning, heating and cooling load calculation. Energy Build 1991;16(1–2):625–34.

[26] Lawrence MG. The relationship between relative humidity and the dew point temperature in moist air: a simple conversion and applications. Bull Am Meteorolgical Soc 2005;86:225–33.

[27] ASHREA, 1997 ASHRAE Handbook. 1997, Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

[28] Anderson DR, Sweeney DJ, Williams TA. Modern Business Statistics with Microsoft Excel. Cincinnati, Ohio; United Kingdom: South-Western; 2003.

[29] Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. Hoboken: John Wiley & Sons; 2006.

[30] IBM, IBM SPSS Statistics 21 Core System User's Guide. 2012, New York: IBM Corporation.

[31] Lam JC, Wan KKW, Lam TNT, Wong SL. An analysis of future building energy use in subtropical Hong Kong. Energy 2010;35(3):1482–90.

[32] Met Office. UKCP09 gridded observation data sets — frequently asked questions. n.d. <http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/faq.html>. [accessed: 4.09.13]

[33] IPCC, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report on the Intergovernmental Panel on Climate Change. 2013, Cambridge and New York: Cambridge University Press.

## Glossary

*CVRMSE:* coefficient of variation of the root-mean-square error

*NMBE:* normalised mean biased error

$R^2$: coefficient of determination

*R:* coefficient of correlation

*RH:* relative humidity (%)

*RMSE:* root-mean-square error

*w:* humidity ratio

$w_s$: saturation mixing ratio

$\beta$: regression coefficient

$\varepsilon$: residual

$\sigma^2$: variance

$\vartheta$: outdoor dry-bulb temperature (°C)