Paper written for submission to Building Research and Information

Assessing the availability and quality of data for key input variables into energy demand forecast models of the UK tertiary sector

Ed Sharp

Paper submitted in part fulfilment of the

Degree of Master of Research in Energy Demand Studies

Energy Institute

University College London

06th September 2011

Title:

Assessing the availability and quality of data for key input variables into energy demand forecast models of the UK tertiary sector.

Author:

Robert Edward Sharp

UCL Energy Institute

Central House

14 Upper Woburn Place

London, WC1H 0NN

ucesres@ucl.ac.uk

Title:

Assessing the availability and quality of data for key input variables into energy demand forecast models of the UK tertiary sector.

Abstract:

The increasing diversity of disciplines utilising demand forecast models has increased the need for accurate projections. Complex models have been developed as a result. These are still dependent on input data however. As a result this study focuses on the identification of appropriate datasets and the assessment of the divergence between sources for the UK tertiary sector.

Six variables used as inputs are identified through the literature and case studies. A literature and web search identifies sources of data on these variables. The availability of data from these data sources is shown graphically. The divergence between the data sources is also calculated and represented as a percentage of the annual mean value.

The divergence between data sources is found to differ across variables due to a number of reasons. Predominantly due to the lack of standard sectoral classification methodology, which is highlighted by a consistent divergence between the values based on NACE and ISIC. Divergence is also shown to be caused by lack of standard calculation methodologies and inconsistent means of harmonisation. These issues alongside a pervasive lack of transparency and inconsistent data storage and transfer methods has led to the study concluding that a centralised data repository is necessary.

Keywords: energy demand, forecast, models, data, data quality, non-domestic, tertiary, demand drivers, data availability.

Word Count: 4,312

Introduction

Energy demand forecast models were incepted as a result of the 1973-74 oil crisis and the subsequent realisation that in order to secure supply potential demand must be projected. Since then they have been used predominantly to ensure that supply meets demand on a national scale. More contemporary applications include those for policy makers: targeting tariffs and subsidies, contributing to climate change models and identifying areas for investment. Models are also used as commercial tools for energy suppliers and others, for example as a means of targeting regions for construction and investment.

Traditionally forecasters have been able to overestimate demand as the negative consequences of excess supply have been limited by cheap fuel and viable short notice generation technologies, predominantly combined cycle gas turbines. Increasing fuel prices due to finite resources and a growing consensus on the need for a reduction of carbon dioxide emissions from supply technologies, alongside the negative economic consequences of over supply have increased the need for accurate forecasting. This paper isolates the recurring variables which forecast models use as inputs, identifies data sources, quantifies the divergence between them and discusses the possible causes of it.

Modelling Methodologies

The need for increased accuracy has led to the iterative development of methodologies by practitioners, complete reviews of the field have been published, examples include Bhattacharyya and Timilsina (2009) and Mehra (2000) which are summarised briefly below with specific examples noted. The first iteration of the models used trends and time series to extrapolate future demand. These have been largely discarded due to

their inability to incorporate societal and technological development and extreme economic events. The next iteration includes end use models: also referred to bottom up or engineering economy which take into account structural technological developments, the models cannot however account for demographic or socioeconomic changes (Mehra, 2000). Decomposition models: which separate energy use into its constituent parts, allowing separate assessment of the factors affecting energy use, these models also struggle to incorporate societal development. The most widely used are econometric models: where demand is depicted as being driven by a series of socioeconomic determinants. Causal relationships have been developed between demand and these determinants through regression analysis; these do not however effectively incorporate structural and technological changes (Lapillonne, 1978) and cannot integrate policy measures and economic shocks (Mehra, 2000). The constituents of this iteration have been combined in what is considered by Bhattacharyya and Timilsina (2009) to be the most significant development: hybrid models which aim to use the most successful aspects of other techniques, examples include all combinations of those outlined here. The most recent and advanced iteration utilises increased computing power and artificial intelligence. These include Neural Networks (see Kermanshahi and Iwamiya, 2002; Taradar Heque and Kashtiban, 2005), Wavelet Networks (Khoa et al., 2004; Ghods and Khalantar, 2008), Genetic Algorithms (El-Naggar and Al-Rumaih, 2005) among others.

Common aspects of models

Despite the diversity of techniques utilised by current practitioners it has been asserted that there are similarities which persist throughout the models. In the context of this study the most important is that all models are considered to be highly data intensive (Mehra, 2000). Mehra adds to this by stating that the dependency of models on data

goes beyond this and as the amount of data increases so does the accuracy of the model. Whilst it is certain that this is not the only factor affecting the accuracy of the model it is probable that those with large amounts of good quality data benefit from it. A lack of quality data has been quantified by Craig et al. (2002) as being responsible for errors of up to 100%. The literature also identified techniques which are dependent on the availability of time series data including Econometric (Bhattacharyya and Timilsina, 2009) and Trend methods (Mehra, 2000).

The tertiary sector

This summary applies to all energy consuming sectors, however the focus of this study, dictated by the research partners Electricité De France (EDF), is a constituent of the non-domestic sector: the tertiary. The non-domestic sector is defined most clearly as that which contains everything that is not used as a home. It is generally regarded as being more heterogeneous than the domestic (Bruhns, 2000; FitzGerald et al., 2002) with a greater diversity of building types and ownership and increased interaction between users and building fabric all complicated by mixed building use being prevalent. The sector is also considered to be relatively deficient in data when compared with the domestic (Denton et al., 2000). Attempts to classify the sector have been made by academics in for example the CaRB project (see Bruhns, 2000) and professionals in the most widely used schemes Eurostat's Nomenclature statistique des Activités économiques dans la Communité Européenne (NACE) (Eurostat, 2008) and the United Nations International Standard Industrial Classifications (ISIC) (UNIDO, 2010). The latter two of these include a distinct tertiary sector the constituents of which are shown in Table 1. These two lists show that definitions do not align and semantic differences exist at all levels. Of note is the inclusion of agriculture, forestry and fishery in ISIC but not NACE and ISIC's definition of the sector as the commercial.

Methods

In order to identify variables on which to collect data, three sources were consulted. The primary source was provided by the research partners EDF in the form of both access to their models of demand in the tertiary sector and the associated list of inputs. EDF are able to utilise their considerable resources to fill data gaps using both modelling and expert knowledge. As a consequence the variables of interest were the ones which they identify as being based on data sources. These were verified using two areas of literature, the first being that which summarises the current state of demand forecasting, outlined briefly in the previous section, verifying that the same variables were used by all forms of models. The second more general literature summarising the factors which affect energy use in non-domestic buildings including: Schipper et al. (1986), IEA (1997), Schipper et al. (2001) and Berry and Jackson (2001). This resulted in a set of common variables which are analysed below.

The literature detailing the availability of data on these variables is limited in scope and depth. Of note are the studies by Macmillan and Köhler (2004) which identifies possible data sources without assessing them and Ó Broin (2007) which assesses the format and content without analysing the difference between the data sources. This study builds on the sources previously identified, firstly through identification of data sources predominantly via a comprehensive web search. The internet has been identified as the predominant source of data by the above literature and EDF. This resulted in a list of data sources including data collation agencies, national statistics organisations and academic and professional research projects. These were analysed individually as part of the wider research project but are not summarised here, more detail can be found in Sharp (2011).

Where available data on each of the variables was downloaded. Figures 1 to 6 show the length of time series available for each of the variables from different sources. As it has been stated in the literature that some model types require data spanning 20-30 time periods for accurate forecasting (Mehra, 2000) an attempt was made to locate data between 1980 and 2009.

Once the data had been downloaded the sources were then compared against each other. In order for the data to be compared considerable cleaning and harmonisation was necessary, where this has been a potential influence on the divergence between datasets details are given in the results section below. The results of the comparison are shown in Figures 1 to 6. These figures also show the maximum range between the data sources. This is expressed as a percentage of the mean so that the difference between the data sources can be compared between variables. Where there are multiple datasets for the whole time series indicative ranges have been given every five years. These ranges facilitate international comparison, detailed in Sharp (2011). Finally the potential reasons for the divergence between the data sources were explored by examining the difference in collection methodologies and classification and artefacts of harmonisation. This investigation was dependent on the availability of metadata from the data sources and relied on transparency. Comments on this have been made in the discussion section of this study.

Results

The following section describes the data that is available for each of the variables identified by the review. The headings represent the list of these variables. A graph representing the available data sources, the time series available from them and

magnitude of divergence is provided for each variable (Figures 1 to 6). The range is quantified as a percentage of the mean annual value so that the divergence can be compared between variables and years. The subsequent discussion section contains analysis of the reasons for the divergence.

Floor space in the sector

Figure 1 summarises the availability of data for this variable. It is clear that there is an overall lack of data with four separate sources. Of these only Odysee provides a time series. There is no data available between 1980 and 1993. The lack of time series provides limited opportunities to assess the magnitude of divergence. In the three instances of more than one dataset for a year the range is variable but small.

Energy consumption in the sector

Figure 2 summarises the availability of data for this variable. There is a large range of data sources each with a significant time series, including data that goes back to 1980. The maximum range does not exceed 50% of the mean for any single year and appears to have halved between 1990 and 2009 after increasing slightly in the decade preceding this. The data provided by national statistics is consistently higher than the rest. The data provided by Odysee closely matches this up to 1996 until steadily reducing to match that provided by the European Environment Agency (EEA), Eurostat and International Energy Agency (IEA).

Population

Figure 3 summarises the availability of the data for this variable. There is a large range of data sources all of which span the entire analysed time period. The maximum range is extremely small up to the year 2000 rising slightly in the ensuing decade, remaining below 1% of the mean.

Employee numbers in the sector

Figure 4 summarises the availability of the data for this variable. There are few available data sources. Those available however provide data for long time series. The maximum range is at a peak of 32% of the mean in 1980 and steadily reduces annually to 12% in 2009. The national statistics dataset and data provided by the United Nations Economic Commission for Europe (UNECE) are consistently the largest of the four. The Odyssee dataset is consistently the lowest. The Organisation for Economic Co-operation and Development (OECD) lies in between until 1990 when it converges with Odyssee and remains at similar levels from the mid 1990's.

Gross Domestic Product (GDP)

Figure 5 summarises the availability of data for this variable. There is a large range of available data sources all of which span the entire analysed time period. The maximum range is 195% of the mean in 1980 which steadily increases to 224% in 2009. The highest values, provided by IEA are clearly wrong however, as they are an order of magnitude higher and at a level which would make British people excessively rich. If excluded from the analysis the remaining data sources align much more closely with agreement between datasets consistent other than values provided by the World Bank between 1998 and 2004. The legend on Figure 5 describes the different methods of calculation used by organisations for GDP.

Gross Value Added (GVA)

Figure 6 summarises the availability of data for this variable. There are few available data sources with varying lengths of time series. Where there are multiple datasets the World Bank gives considerably higher values (up to 160% of the mean). Odysee provides

the lowest values whilst the limited time series given by national statistics (1996 – 2000) lies in between.

Discussion

Floor space in the sector

This variable has been identified in the literature as the key driver of energy demand in the non-domestic sector (Zhou and Lin, 2007). Therefore the significance of the lack of data is high. The lack of time series is also important as it negates the use of the data in models which are dependent on them.

Although the lack of data negates in depth analysis of divergence between data sources the available data does highlight two important points. The first of which is the impact of differing sectoral definitions. The larger of the two data points in 1993 represents the floor area of the entire non domestic stock as described in the introduction found by the CaRB project (Figure 1). The smaller value represents the floor area of the tertiary sector given by Odyssee based on the NACE classification described in Table 1. Although the constituents of the non-domestic sector not included in NACE's definition of the tertiary should increase the value it is likely that it would be more than is seen in this analysis. The data points in 2003 demonstrate the same problem where Ecoheatcool's value claims to represent all commercial and public buildings which would be expected to be significantly higher than the value provided by Odysee, however the difference is small.

The second issue highlighted by this data is the lack of transparency and consequent difficulties in tracing the source. The associated metadata for the data sources shown in Figure 1 is reasonably good, especially CaRB which is unique in that it has a large amount

of associated literature. However there are still problems, of note is that Ecoheatcool state national statistics as a source of data despite the variable not being directly available from them. It is difficult therefore to take this data as accurate or authentic.

Energy consumption in the sector

Figure 2 demonstrates that the data provided by national statistics is consistently higher than what appears to be a consensus between the majority of the remaining data sources. This divergence may be due to the varying definition of energy used between the sources. This could be either primary energy which represents the amount contained in the fuel used to produce it or delivered energy representing that used at the end of the cycle accounting for production and transmission losses. This is however very rarely specified in the metadata, negating any attribution of error and highlighting the need for increased depth and detail in associate information.

What can be analysed however is the divergence associated with the classification of the sector. All of the data sources providing the lower values use NACE as described in Table 1 whereas national statistics values are based on ISIC. The difference is therefore most likely due to the inclusion of agriculture, forestry and fishery in the latter. The reducing range is likely due to the contracting size of the sectors or reducing energy intensity.

Population

The small levels of divergence shown in Figure 3 demonstrate the benefit of established collection mechanisms. Population is systematically recorded via birth and death certification, period census collection (every 10 years) and border control. Despite this there are accepted sources of error including illegal immigration, unrecorded births and deaths and inconsistent inclusion of expatriates.

The divergence between the sources for this variable may also be due to the level of precision used when storing data. National statistics, Odysee and OECD give population to the nearest one thousand people. Eurostat, the World Bank and UNECE give population to the nearest single person and IEA to the nearest one hundred thousand. Using the values provided for 2009 these differences in precision create a potential range of 0.06% of the mean between the largest and smallest possible value. As Figure 3 shows the actual difference is 0.43% of the mean suggesting that this is not the only source of error. However for the data before the year 2000 the ranges are of this magnitude suggesting differences in the precision of data storage can be considered to be the cause of the majority of the difference.

Employee numbers in the sector

The data summarised in Figure 4 appears to display the same artefacts of divergent classification schemes shown in the analysis of energy consumption of the sector (Figure 2). The largest of the datasets (national statistics) follows the ISIC classification and the lowest (Odysee and OECD) follows NACE. Therefore the inclusion of agriculture, forestry and fishery in the former is the likely reason for higher values. As with energy consumption the values converge gradually over the time series and the difference roughly halves between 1990 and 2009 giving further impetus to the assertion that there is an artefact of the difference being due to a shrinking sector.

The lack of transparency and complete metadata is highlighted by the fact that the classification methodology is not stated by UNECE. It is likely that ISIC is used as the values closely match those provided by national statistics however this would not be apparent if the data was viewed in isolation.

Gross Domestic Product (GDP)

The data provided by IEA shown in Figure 5 is, as stated above, clearly incorrect. The reasons for this are unclear although it is likely that a conversion factor has been applied at some point which has caused the value to be multiplied by approximately 5. It is possible that this is due to the fact that the values are given as per capita and the wrong population data has been used. This again highlights the lack of transparency as the statistics used by data sources to give per capita value are not explicitly stated. However as the analysis of the population variable shows the impact of using different datasets should be minimal as the divergence is small between sources. Therefore it is likely that the error has been introduced in another way.

This variable also demonstrates how data can be complicated by complex and differing methods of calculation. The legend on Figure 5 shows that there are different ways of converting between currencies to harmonise data into an internationally comparable form (most commonly the United States of America Dollar (USD)). These represent possible causes of divergence and are particularly problematic if not explicitly stated.

This metric benefits from being used by a range of disciplines and therefore collected as standard by many agencies. As a consequence it is part of the European System of National Accounts (ESA95) and is collected using defined methodologies by all European Union countries. The impact of this can be seen in the minimal divergence between the datasets in Figure 5 excluding those provided by IEA.

Gross Value Added (GVA)

Unlike GDP, GVA is not collected as standard in any pan European schemes which is reflected in the lack of data shown in Figure 6. Also unlike GDP it only uses one method of calculation. The expected result of which would be closely aligned data between sources. This is not the case however, which is likely to be a result of the metric relying

on consistent methods of classifying the sector which it represents. Like the other variables dependent on this (energy consumed and employee numbers) the data source which use ISIC (national statistics) gives higher values than Odysee using NACE. In this case however there is another data source which is significantly higher than both. The World Banks metadata shows that it uses ISIC but describes the target sector as "services etc." This includes transport and governmental services on top of the ISIC sub sectors shown in Table 1. This certainly accounts for some of the increased values. The extent of the contribution is unclear.

Conclusions

The review of the literature has demonstrated that the accuracy of demand forecast models is dependent at least in part on the availability and quality of data over lengthy time periods (Mehra, 2000; Craig et al., 2002). It has also shown that there is a relative lack of this data for the focus of the research: the non-domestic sector (Denton et al., 2000). An analysis of case study models provided by EDF and associated demand forecasting literature has identified 6 variables which are drivers of energy demand and commonly used in models as input. A literature and web search has identified sources of data on these variables. The availability of data from these data sources is shown in Figures 1 to 6.

The divergence between the data sources has been attributed to a number of different reasons that in some cases are pervasive across multiple variables. The clearest divergence has been a result of the use of different schemes for classifying the sector of interest: the tertiary. At the highest level there are semantic differences with ISIC referring to it as tertiary, NACE as commercial and the World Bank as services etc.

Below this ISIC has included agriculture, forestry and fishery resulting in higher values for variables dependent on the classification of the sector (energy consumption, employee numbers and GVA) than NACE which has excluded these sub sectors. Data sources have also been shown to calculate metrics in different ways, in particular GDP. Finally there have been artefacts of harmonisation again evident in the per capita GDP values.

These are issues that could be solved or at least made more transparent with the existence of a centralised repository for this type of data. This could serve to ensure that metadata is complete and consistent, ensuring that the reasons for differences between datasets are clear. The potential implications for the models that use the data could then be anticipated.

Alongside issues with classification and harmonisation there are other ways in which the data varies between data sources. The format in which data is stored is inconsistent. Most commonly used are spreadsheets with a small number of interoperable files (XML) with fewer still web based tables. This necessitates considerable formatting in order to make this data usable. The precision of the data that is stored in these files varies as shown in the example of population statistics above. These differences can be mistaken for accuracy unless completely transparent. These issues could also be solved through a repository. In its place however it is possible to avoid many of these issues or at least make the consequences clear though two mechanisms. Primarily the use of standard classification schemes, followed by comprehensive metadata. The successful application of these types of changes and potential resultant increases in model accuracy will increase their effectiveness. This has positive implication for all of the practitioners of demand forecasting in the increasing variety of fields of application.

Future Work

This research is part of a wider project which has assessed the same variables and data sources for all 27 European Union Countries (Sharp, 2011). This data can be disseminated in the same way as here. As described above an aim at the end of this project is to investigate the possibility of creating an online repository which contains the data analysed by this project. This repository would aim to create a "one stop shop" for the data increasing ease of access and ensuring consistent standards of storage, harmonisation and classification where possible. The repository would also ensure consistent and high quality metadata ensuring transparency of collection methodologies and identification of potential reasons for divergence.

Possible extensions of this work include measuring the propagation of error through existing demand forecast models for each of the variables quantifying the final implications of divergent data on demand forecasts. Comparing the data to that which is available for the domestic sector to assess the validity of the claim that it has larger volumes of good quality data associated with it than that analysed in this study (Denton et al., 2000). Also of interest is whether the claim that economically more developed countries have better quality associated data (Bhattacharyya and Timilsina, 2009), which can be tested using the data collected for this project between European countries. The amount of data collected also means that it would be possible to identify other drivers of energy demand by analysing possible reasons for sudden increases and decreases. These may include political, economic and climatic drivers.

Acknowledgements

This research was undertaken with the collaboration of EDF. EDF provided access to demand forecast models and advice on their structure and operation. EDF also provided access to the Odysee database. All other data is accessible for no charge.

References

Berry, A. and Jackson, T. (2008). Social and institutional trends affecting non-domestic sector energy consumption: A preliminary scoping study. *Resolve Working Paper*. Guildford, England.

Bhattacharyya, S. and Timilsina, G. (2009). Energy Demand Models for Policy Formulation: A comparative study of energy demand models. *Policy research working paper 4866*. The World Bank.

Bruhns, H. (2000). Property taxation data for non-domestic buildings in England and Wales. *Environment and Planning B: Planning and Design*, 27(1), pp. 33 – 49.

Craig, P. Gadgil, A. and Koomey, J. (2002). What can history teach us? A retrospective examination of long-term energy forecasts for the United States. *Annual Review of Energy and the Environment*, 27, pp. 83 - 118.

Denton, F. Mountain, D and Spencer, B. (2000). A model of energy demand in the US commercial sector with declining rate schedule. *Research Report 346*. Canada, McMaster University.

DUKES. (2010). *Digest of United Kingdom energy statistics*. UK: DECC. Available from: http://www.decc.gov.uk/en/content/cms/statistics/publications/dukes/dukes.aspx. Last Accessed: 18th May 2011.

El-Naggar, K. and Al-Rumaih, K. (2005). Electric load forecasting using genetic based algorithm, optimal filter estimator and least error square technique: Comparative study. *Transactions of Engineering, Computing and Technology*, 6: pp. 138 - 142.

Eurostat. (2008). Eurostat Yearbook 2008. Available From: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CD-07-001-01/EN/KS-CD-07-001-01-EN.PDF. Last Accessed: 28th July 2011.

Fitz Gerald, J. Hore, J. and Kearney, I. (2002). A model for forecasting energy demand and greenhouse gas emissions in Ireland. *Working paper 146*. The Economic and Social Research Institute (ESRI).

Ghods, L. and Khalantar, M. (2008). Methods for long term electric load demand forecasting; a comprehensive investigation. In: *Proceedings of the IEEE International Conference on Industrial Technology, April 2007.* Chengdu, China: ICIT.

IEA. (1997). *Indicators of energy use and efficiency: Understanding the link between energy and human activity*. Paris: Organisation for Economic Co-operation and Development (OECD).

Kermanshahi, B. and Iwamiya, H. (2002). Up to year 2020 load forecasting using neural nets. *Electric power system research*, 24 (9), pp. 789 - 797.

Khoa, T. Phuong, L. Binh, P. and Lien, N. (2004). Power load forecasting algorithm based on wavelet packet analysis. In: *Proceedings of The International Conference on Power System Technology*, 21-24 *November 2004*. Singapore: POWERCON, pp. 987 - 990.

Lapillonne, B. (1978*). MEDEE 2: A model for long-term energy demand evaluation*. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Macmillan, S. and Köhler, J. (2004). Modelling energy use in the global building stock: a pilot survey to identify available data sources. *Technical Report 6*. Tyndall Centre for Climate Change Research.

Mehra, M. (2000). Demand forecasting for electricity. *International conference on regulation and governance, November 2000.*

Ó Broin, E. (2007). *Energy Demand of European Buildings: A mapping of available data, indicators and models*. Unpublished Thesis (MSc), Chalmers University of Technology.

Saunders, M. Lewis, P. and Thornhill, A. (2003). *Research methods for business students.* Harrow: Pearson Education Limited.

Schipper, L. Meyers, S. and Ketoff, A. (1986). Energy use in the service sector: An international perspective. *Energy policy,* 14 (3), pp. 201–218.

Sharp, R. (2011). Assessing the suitability of data sources for input into European energy demand forecast models. Unpublished Thesis (MRes), University College London.

Taradar Heque, M. and Kashtiban, A. (2005). Application of neural networks in power systems; A review. *Transaction of Engineering, Computing and Technology*, 6, pp. 53 - 57.

UNIDO. (2010). Industrial Statistics Database at the 3 and 4 digit level of ISIC Code (Revision 3). Available From: http://www.esds.ac.uk/international/support/user_guides/unido/InSt3x-UserGuide.pdf. Last Accessed 20th July 2011.

Zhou, N. and Lin, J. (2007). The reality and future scenarios of commercial and building energy consumption in China. *Energy and Buildings,* 40 (12), pp. 2121 – 2127.

Table of Figures

Table of Tables

Figures

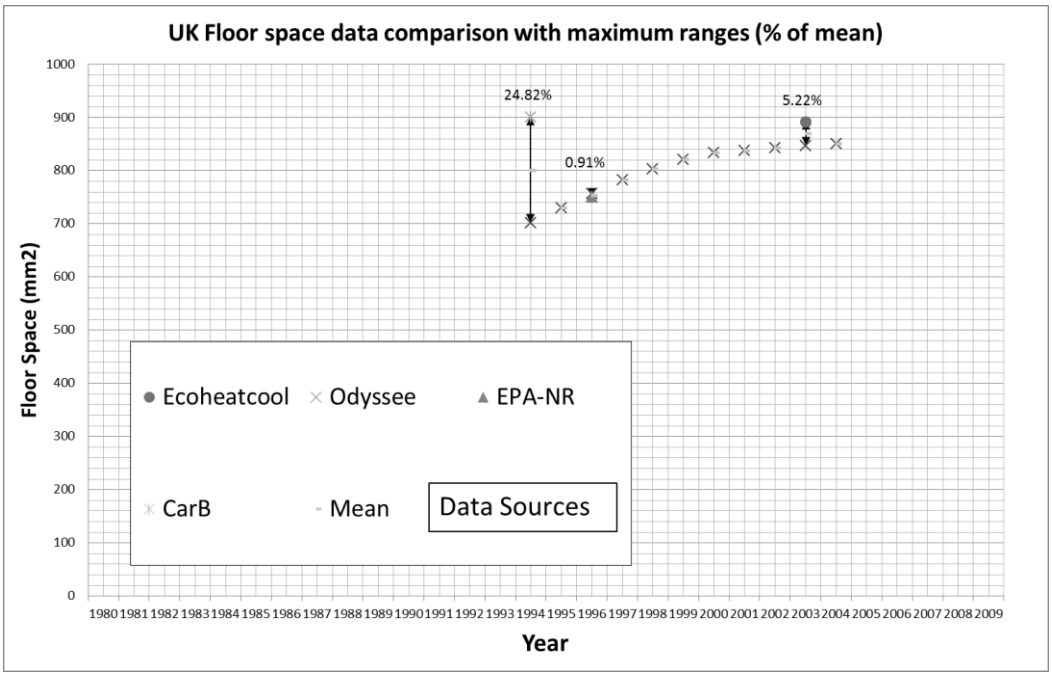

Figure 1 - Floor space data for the UK, 1980 – 2009, including maximum annual range as a percentage of mean.

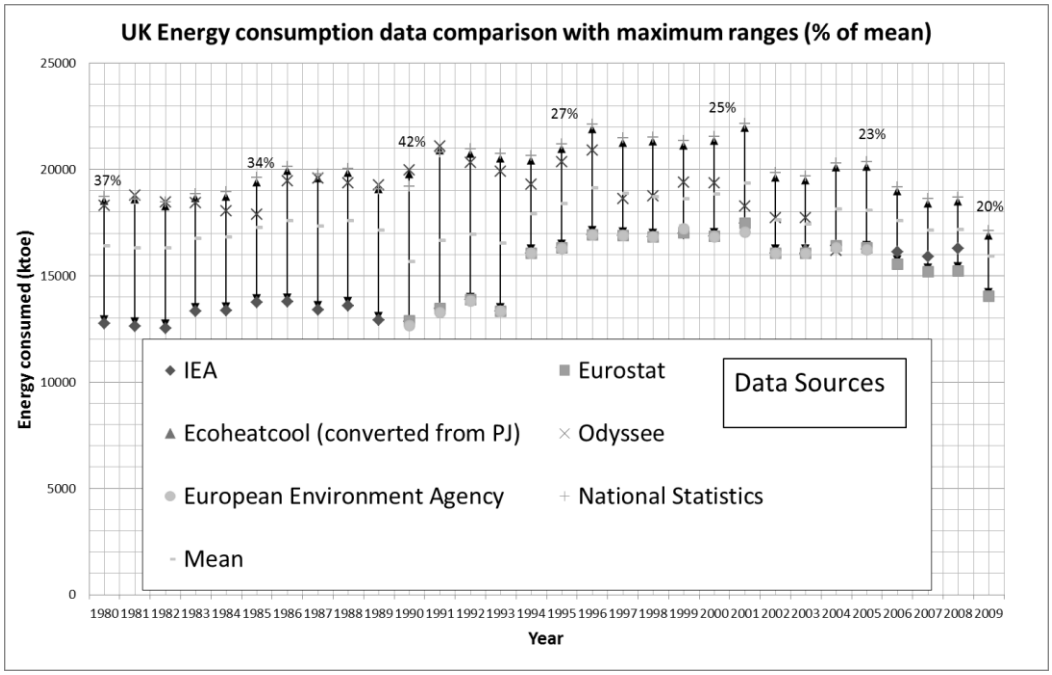

Figure 2- Energy consumption data for the UK, 1980 – 2009, including maximum annual range as a percentage of mean.

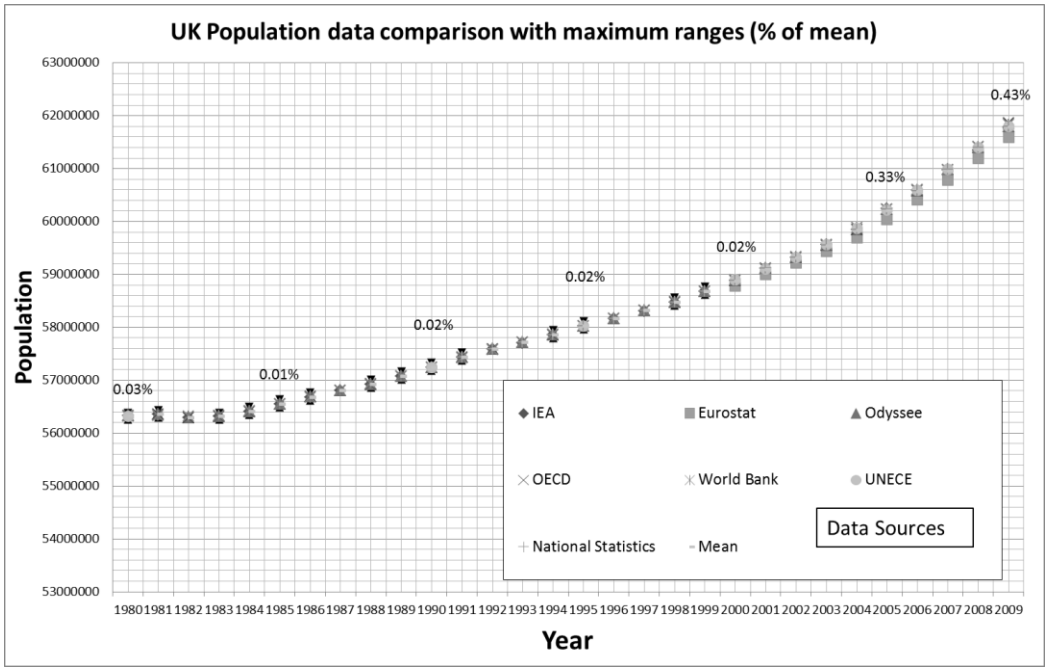

Figure 3 - Population data for the UK, 1980 - 2009, including maximum annual range as a percentage of mean.

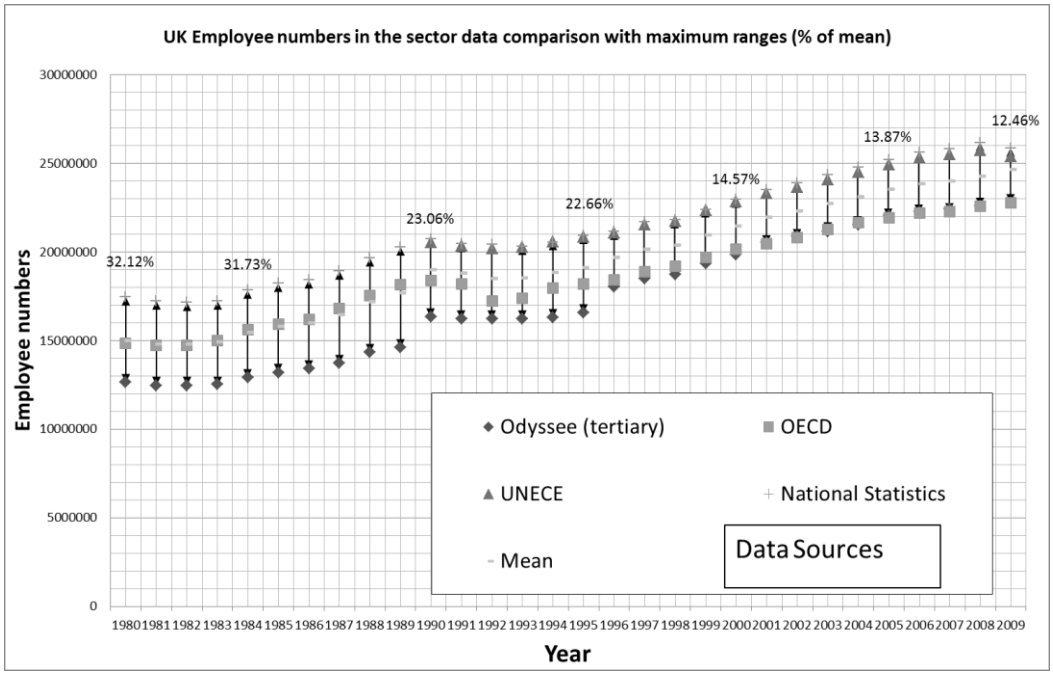

Figure 4 - Employee numbers in the sector data for the UK, 1980 - 2009, including maximum annual range as a percentage of mean.

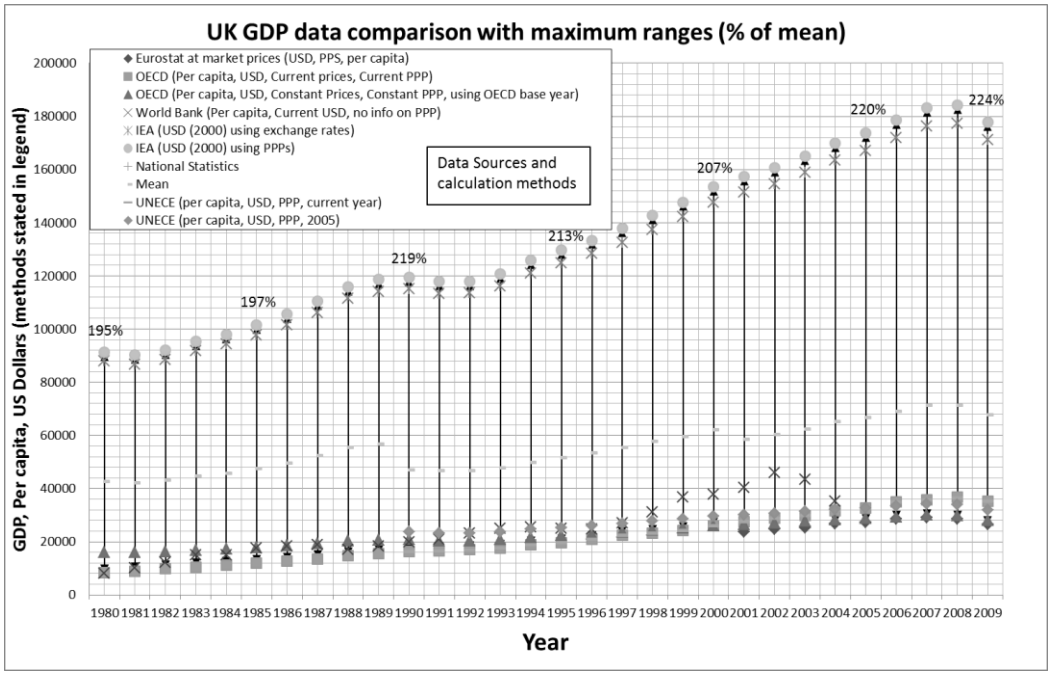

Figure 5 - GDP data for the UK, 1980 - 2009, including maximum annual range as a percentage of mean.

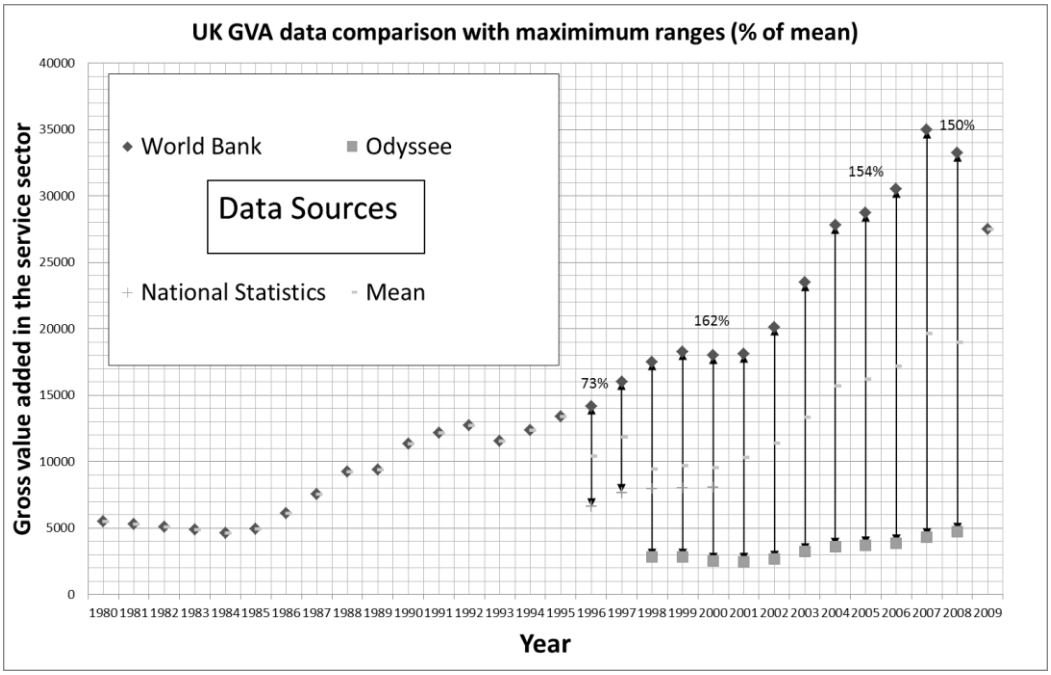

Figure 6 - GVA data for the UK, 1980 - 2009, including maximum annual range as a percentage of mean.

Tables

NACE (Tertiary)	ISIC (Commercial)
Wholesale & Retail Trade; repair of motor vehicles and motorcycles	Wholesale and Retail Trade; Repair of Motor Vehicles, Motorcycles and Personal and Household Goods
Accommodation and food service activities	Hotels and Restaurants
Financial, insurance and real estate activities	Real Estate, Renting and Business Activities
Administrative and support service activities	Post and telecommunication, Financial Intermediation
Education	Education
Human health and social work activities	Health
Other NACE activities	Miscellaneous
	Public administration and defence
	Agriculture, Forestry and Fishery (as separate sub sectors

Table 1 - Constituents of the NACE tertiary and the ISIC commercial classification schemes.