Technical Papers supporting SAP 2009

A review of the relationship between floor area and occupancy in SAP

Reference no.	STP 09/NOFA01
Date last amended	04 April 2008
Date originated	04 April 2008
Author(s)	John Henderson, BRE

Introduction

SAP and BREDEM use an equation to estimate the number of occupants from the floor area of a dwelling. This is always used in SAP, but is optional in BREDEM (the actual number of occupants should be entered instead where it is known). An earlier BRE internal study using English House Condition Survey (EHCS) data for 2001 suggested that the existing relationship gives poor approximations for very large homes.

This brief note describes how EHCS data has been used to develop an improved equation to use in future versions of SAP and BREDEM.

Analysis

Data showing 'total floor area', 'number of occupants' and the 'grossing factor' for all the 'households' surveyed as part of the EHCS studies of 2002, 2003, 2004 and 2005 was inserted into a spreadsheet, giving records for a total of 32009 homes. It is useful to define some of the terms used here:

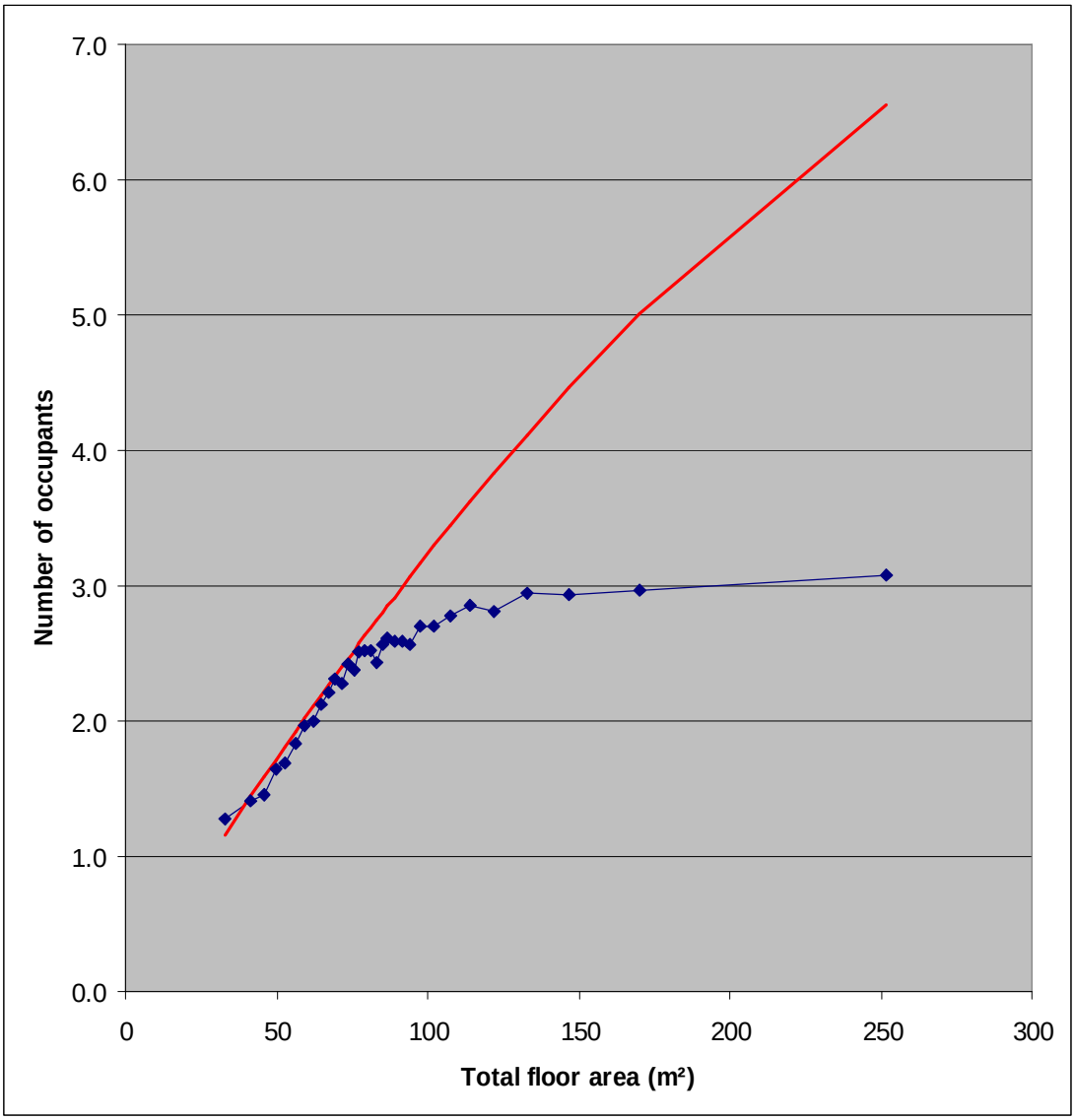
- 'Households' means occupied dwellings. Some unoccupied dwellings were surveyed as part of the EHCS, but these were excluded because this analysis was specifically to consider how many occupants a dwelling is likely to contain based on its floor area. (The EHCS term for the group which includes unoccupied homes is 'dwellings')
- 'Total floor area' (TFA) includes the entire internal floor plan of all storeys of the dwelling. This means for example that the area above the stairs (which is not in fact 'usable' floor space) is included, as is the inaccessible floor area under internal walls. This is appropriate for this exercise because the floor areas used as inputs to SAP and BREDEM are defined in the same way.
- 'Number of occupants' (Occ) is the count of the total number of people (adults and children) living permanently in the dwelling.
- The 'grossing factor' is a weighting that should be applied to the data for each household when performing any kind of averaging. Each house represents between a few hundred and several thousand dwellings, such that the total represented is equal to the total number of households in England[1].

It is not very useful to start with a simple plot of occupancy against floor area for all the homes because of the immensely large scatter in the data, though this is plotted later in this note. In order to reveal the trend of interest therefore, the data was sorted by the floor area of the dwellings then grouped as follows:

- The average floor area and occupancy of the first (i.e. smallest) 1000 households was calculated (weighted using the grossing factors)
- This was repeated for the next group of 1000 homes and then each subsequent group
- The final group (containing the largest homes) was slightly enlarged to include the 'spare' 9 homes, resulting in a group size of 1009

The resulting data set therefore contained 32 values for each parameter. By then plotting the derived occupancy figures against floor area, the originally hidden trend became apparent.

[1] But note that in the whole sample, since 4 years of data were used in combination, the total number of dwellings represented by the data was in fact around 4 times the number in England.



As had been found in the earlier analysis, an occupancy level appearing to saturate at around 3 in large homes is indicated. For the smallest homes the number of occupants appears to tend towards 1. This contrasts sharply with the existing SAP and BREDEM relationship (shown in red) which suggests that the number of occupants continues to rise far beyond 3 in homes with floor areas of greater than 90m². (Though not shown on the graph, this eventually saturates at around 9 for the largest homes.) The existing relationship also gives values of less than 1 occupant for very small homes (less than 29m²)[2].

As a proposed better fit, the following equation, based on an s-curve, was developed:

[2] This sounds so small it is trivial, but 200 homes out of the 32009 (0.6%) recorded in the EHCS data were reported to have a floor area of less then 29m², so these should be catered for.

For TFA >= 13.9

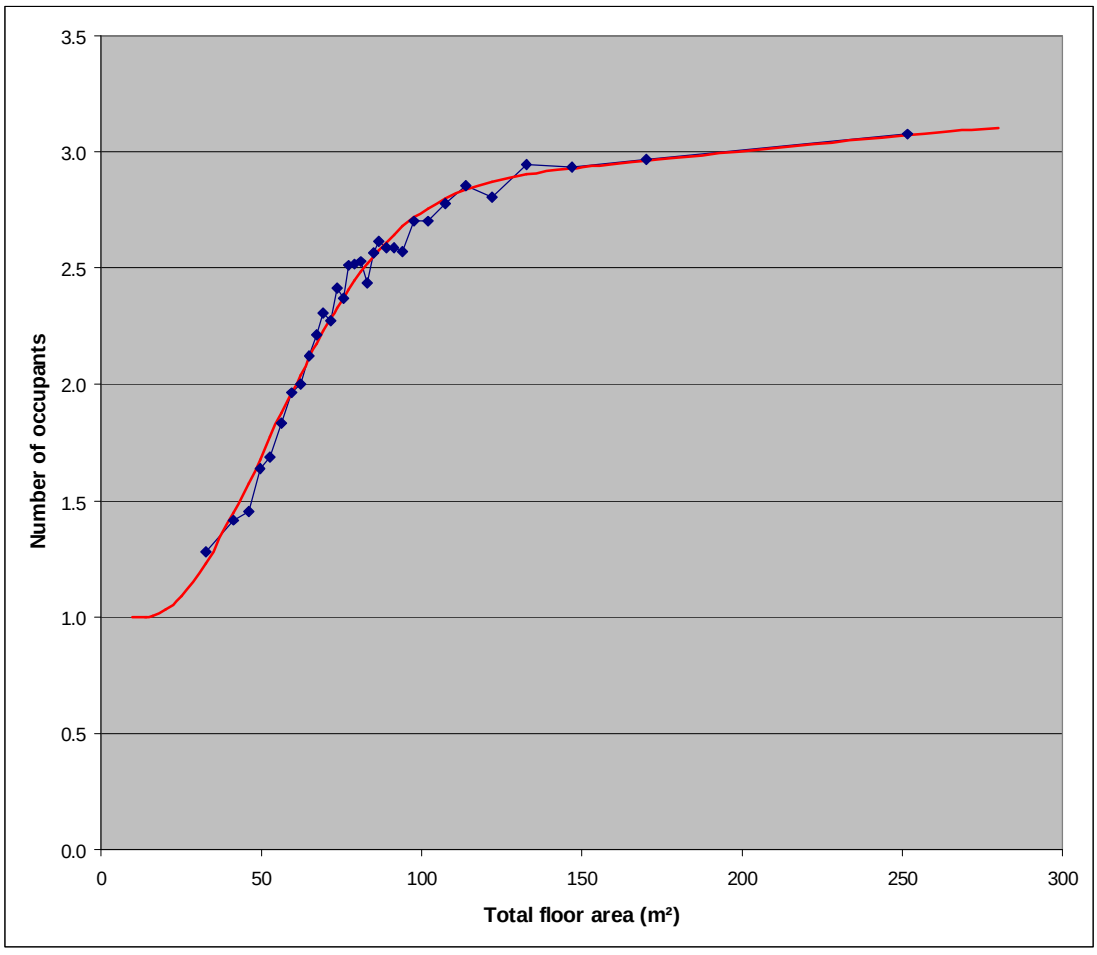
Occ = 1 + 1.76 (1-exp (-0.000349 (TFA-13.9)^2)) + 0.0013 (TFA-13.9)

For TFA < 13.9

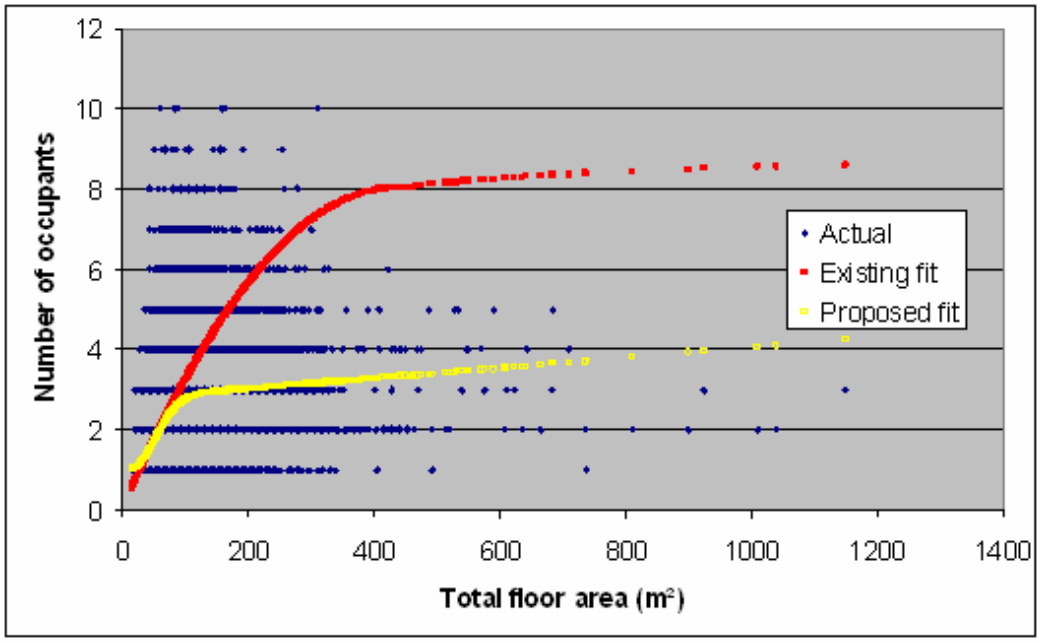
Occ = 1

(3 significant figures for each constant is an adequate level of precision.)

The following graph shows the good fit to the data this achieves.



To check that the grouping method used was not leading falsely to the conclusion that the existing fit is inferior to that proposed, the original full data set was returned to. The existing SAP and BREDEM formula was used to predict the number of occupants. Alongside this, the predicted number of occupants was calculated using the proposed new formula. The following graph shows these plotted upon the full data set.



Since it is hard to see through the mass of points which fit is best, the following tests were carried out on each fit to confirm this.

- The sum of the squares of the differences between the predicted and actual results was calculated
- The percentage of occasions on which the predicted result was within 1 of the actual result was calculated
- The percentage of occasions on which the predicted result was within 2 of the actual result was calculated

The results of this showed that the new fit was better than the existing one by each measure. The figures obtained are summarised in the following table. This analysis was also carried out for homes with floor areas of greater than 100m² and greater than 200m² for comparison.

	All homes			**Homes >100m²**			**Homes >200m²**		
Fit used	Sum of errors²	Within 1	Within 2	Sum of errors²	Within 1	Within 2	Sum of errors²	Within 1	Within 2
Existing	71482	53%	85%	37064	29%	62%	11675	6%	18%
Proposed	49985	60%	92%	15972	51%	87%	1562	40%	88%

The modesty of the improvement for all homes is because the large majority of homes fall between 50 and 100m², where the existing equation gives a reasonable fit. This is presumably why it has taken so long for the poorness of fit for larger homes to be noticed. The improvement is much more marked when only large homes are considered.

Conclusion

The equation proposed in this note gives a better approximation to the actual number of occupants than the existing equation used in SAP and BREDEM. It is therefore suggested that it is used in future versions as the default method of calculating the number of occupants. This should give a superior estimate compared to the existing equations for most homes, and especially for very large ones. It should always be kept in mind however what is implied by the very large scatter exhibited by the full ungrouped data set: occupancy varies greatly for a given floor area. In other words, floor area is not a particularly good indicator of the number of occupants. Thus it is always preferable to use the actual number of occupants in a BREDEM calculation where that is known. Of course that would not be appropriate for SAP where standardised occupancy is desirable.

The implication of making this change is that in large homes the amount of energy predicted for hot water use, lighting and appliances will fall significantly (in the absence of other alterations to their calculation procedures), since these are all functions of occupancy. Gains from these sources will therefore also fall. Consequently space heating will tend to go up in absolute terms and especially in relative terms as a proportion of total energy use.