Journal of Environmental Radioactivity xxx (2011) 1–7

Contents lists available at ScienceDirect

Journal of Environmental Radioactivity

journal homepage: www.elsevier.com/locate/jenvrad

Two error components model for measurement error: application to radon in homes

Nezahat Hunter*, Colin R. Muirhead, Jon C.H. Miles

Health Protection Agency, CRCE (Centre for Radiation, Chemical and Environmental Hazards), Chilton, Oxon OX11 0RQ, UK

A R T I C L E I N F O

Article history:
Received 28 July 2010
Received in revised form
4 May 2011
Accepted 31 May 2011
Available online xxx

Keywords:
Additive error
Bayesian approach
Measurement error
Multiplicative error
Radon concentrations
Repeated measurements

A B S T R A C T

In this paper, a simple model for analysing variability in radon concentrations in homes is tested. The approach used here involves two error components, representing additive and multiplicative errors, together with variation between-houses. We use a Bayesian approach for our analysis and apply this model to two datasets of repeat radon measurements in homes; one based on 3-month long measurements for which the original measurements were close to the current UK Radon Action Level (200 Bq m^{-3}), and the other based on 6-month measurement data (from regional and national surveys), for which the original measurements cover a wide range of radon concentrations, down to very low levels. The model with two error components provides a better fit to these datasets than does a model based on solely multiplicative errors.

1. Introduction

Radon (^{222}Rn) is a naturally occurring radioactive gas that is produced from uranium (^{238}U) which is found in rocks, soil and building materials. Radon gas can enter a dwelling by the process of diffusion from building materials and from the ground through cracks and other holes in the foundations. These concentrations can vary widely from day to day, but generally follow a pattern of seasonal variation with high levels in the winter and low levels in the summer. If a measurement is made over a period of less than a year (eg. three or six months), then it is necessary to normalise the result to estimate the average concentration during the year. This can be done by using seasonal correction factors (Miles and Howarth, 2000; Wrixon et al., 1988). Since the indoor–outdoor temperature difference is the driving force to the entry of radon into dwellings, correction for the outdoor temperature rather than the season is appropriate (Miles, 1998). This temperature correction can be made by using temperature normalisation factors based on the indoor/outdoor temperature difference observed during a particular measurement (Miles, 1998).

Uncertainties in estimates of long-term mean radon concentrations arise from a number of factors: uncertainties in the detector results themselves, uncertainties in extrapolating from a limited period to a full year, and uncertainties in extrapolating from a single year to the long-term average. In particular, indoor radon concentrations can be highly variable from year to year (Hunter et al., 2004). Other factors known to affect the variability in radon measurements include house specific factors (eg. house type, double glazing and floor type), structural alterations to the building (eg. cavity wall insulation, blocking of a chimney and sealed floors), variation in the weather, variation in the lifestyle of those living in the dwelling and the location of dwellings (particularly the geology of the area) (Hunter et al. 2009, 2004; Gerken et al., 2000; Gunby et al., 1993).

Epidemiological case–control studies have been conducted in various countries worldwide to evaluate the risk on lung cancer and residential radon exposure. Three combined analyses have been published in recent years: a European pooling study, in which individual data from the main European studies of residential radon and lung cancer were brought together and reanalysed (Darby et al., 2005, 2006): a North American pooling study (Krewski et al., 2005, 2006) based on studies conducted in the USA and Canada; and a combined analysis of two studies in China (Lubin et al., 2004). All three pooled analyses found that residential radon exposure was associated with an increased risk of lung cancer. However, random variability (measurement error) in radon measurements is likely to have biased risk estimates towards zero. Moreover, uncertainties in estimates of average annual radon concentrations in individual houses would result in the distribution of estimated radon levels

* Corresponding author. Tel.: +44 1235 822803; fax: +44 1235 833891.
E-mail address: nezahat.hunter@hpa.org.uk (N. Hunter).

0265-931X/$ – see front matter
doi:10.1016/j.jenvrad.2011.05.009

being wider than the true distribution of long-term average radon levels in houses (Darby et al., 2006). This wider distribution may lead to an overestimate of the proportion above the Action Level (200 Bq m^{-3}) in UK unless allowance is made for this. Methods to correct such errors are available and have been taken account in epidemiological studies (Darby et al., 2006; Heid et al., 2006; Lagarde et al., 1997) and radon mapping (Miles et al., 2007).

This study aimed to evaluate measurement uncertainty in estimates of average annual radon concentrations in individual houses. We describe a new multilevel model for radon measurements that allows for three sources of variability. These sources are: variation between-houses, and multiplicative and additive measurement errors. We show how this model can be fitted from a Bayesian approach using the WinBUGS program (Spiegelhalter et al., 1996). This study was performed using two different datasets. For the first, measurements took place over a period of 3 months, while for the second dataset measurements were made over a period of 6 months. Both datasets contain repeat measurements within each of a number of dwellings. However, the two datasets cover differing ranges of radon concentrations. An outdoor temperature correction was used rather than a seasonal correction to adjust radon results for the reasons stated earlier.

2. Data and descriptive analyses

The two different datasets analysed were collected by the National Radiological Protection Board (now part of the Centre for Radiation, Chemical and Environmental Hazards of the Health Protection Agency) throughout the United Kingdom (Kendall et al., 2005; Wrixon et al., 1988). The first dataset contains repeat 3-month measurements in 4566 dwellings and the second comprises repeat six-month measurements in 2777 dwellings. The homes selected all had an initial outdoor temperature corrected radon measurement. The measurements for both datasets were carried out during two consecutive three month or two consecutive six-month periods, using passive radon detectors placed in both the main living area and an occupied bedroom in the house.

2.1. 3-month repeat measurement data

Table 1 presents summary statistics for the 4566 pairs of repeated 3-month radon measurements, both as the uncorrected data and as corrected data adjusted for outdoor temperature. The 3-month dataset was based on homes where the original measurement was close to the current UK radon action level, namely 200 Bq m^{-3}. The arithmetic mean, geometric mean and geometric standard deviation of the measured radon concentrations were 191 Bq m^{-3}, 169 Bq m^{-3} and 1.63 respectively based on the temperature corrected data and were similar to the corresponding values based on the uncorrected data.

Radon concentrations in houses are frequently found to follow a log-normal distribution. Fig. 1 shows a normal probability plot of the temperature corrected natural logged radon concentrations of the 3-month measurements. This plot shows a reasonably linear pattern in the centre of the data (see Fig. 1). However, there are too many values in the tails at each end of the distribution, resulting in

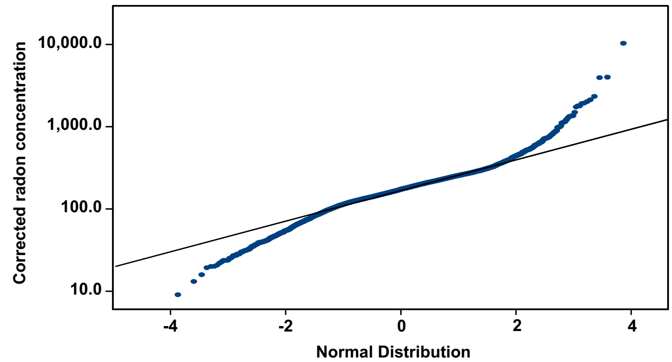

Fig. 1. Normal probability plot of the 3-month repeated radon concentration data (Bq m^{-3}) with correction for outdoor temperature from 4566 houses, under a logarithmic transformation. The solid line represents the linear fit to these data, based on a normal distribution.

departures from the line fitted on the basis of a normal distribution. This may reveal a mixture of two log-normal distributions representing arrange of different geological terrains. Furthermore, a long-tailed data such as this might be fitted better by a long-tailed t-distribution (ref). However, the use of this distribution in place of a log-normal distribution had little influence on the following results.

Fig. 2a shows that there was a statistically strong agreement between the initial and final measurements in the same dwelling (correlation coefficient, $r = 0.58$). In particular, most data points are not close to the line of equality. An alternative, more informative plot - shown in Fig. 2b - is the difference between the logged repeat and initial measurements versus the average of the two logged measurements for each dwelling, which is equivalent to a log standard deviation versus log mean plot. It is apparent from the plot that the difference between the logged radon measurements (R2−R1) increases with increasing values for the mean of the logged measurements, as shown by the positive slope of the regression ($P < 0.001$).

2.2. Six-month repeat measurement data

Table 2 shows summary statistics for the 2777 pairs of repeated six-month radon measurements, both as uncorrected data and corrected for temperature and outdoor radon concentrations. The latter adjustment for outdoor radon was made in order to make the data a more accurate representation of the true values (Gunby et al., 1993). In particular, the mean UK outdoor radon level of 4.1 Bq m^{-3} was subtracted from each measurement before taking the logarithmic transformation. When these adjustments for outdoor radon and temperature were made to the 6-month data, visually from Fig. 3, we obtained a better normal probability plot than was the case with adjustment for temperature alone (see Fig. 3). In contrast, when adjustments were made for outdoor radon in addition to the temperature to the 3-month data, the normal probability plot remained the same as in Fig. 1. This is due to the fact that 3-month measurements were close to the current UK action level (200 Bq m^{-3}) and subtracting outdoor radon concentration didn't

Table 1
Summary statistics for radon concentrations (expressed in Bq m^{-3}) based on 3-month measurements in 4566 houses.

	Radon concentrations					
	Min. (Bq m^{-3})	Max. (Bq m^{-3})	Arithmetic mean (Bq m^{-3})	Std. Deviation (Bq m^{-3})	Geometric mean (Bq m^{-3})	Geometric std. deviation
Uncorrected data	8.29	6217	189	132	170	1.55
Temperature corrected data	9.03	10285	191	165	169	1.63

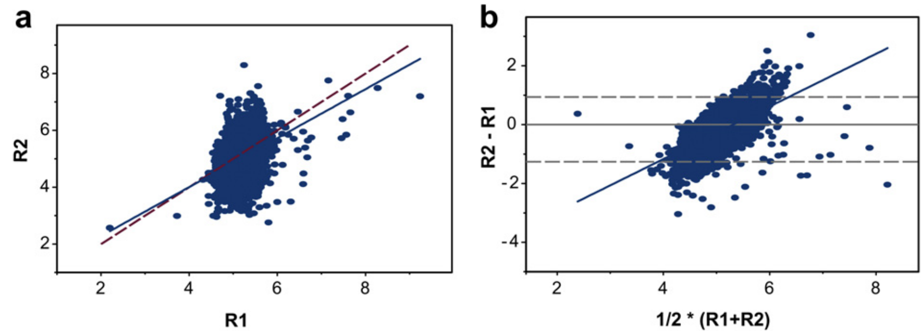

Fig. 2. The relationship between two 3-monthly radon measurements in the same dwelling (data are temperature corrected log radon measurements). (a) R1 (initial measurement) versus R2 (final measurement), with the line of equality (broken line) and a regression line (solid line). (b) Change in logged radon measurements (R2−R1) plotted against the average of the initial and final logged measurements ((R1+R2)/2), with the range corresponding to two standard deviations (horizontal dashed lines) and a regression line (solid line).

Table 2
Summary statistics for radon concentrations expressed in Bq m^{-3}, based on six-month measurements in 2777 houses.

	Radon concentrations					
	Min. (Bq m^{-3})	Max. (Bq m^{-3})	Arithmetic mean (Bq m^{-3})	Std. Deviation (Bq m^{-3})	Geometric mean (Bq m^{-3})	Geometric std. deviation
Uncorrected data	1.56	4014	88	215	27.25	3.88
Temperature corrected data	1.27	3227	90	217	27.61	3.92
[a]Temperature and outdoor radon corrected data	0.01	3222	86	217	19.30	5.44

[a] Outdoor radon values ≤4.1 Bq m^{-3} were set to 4.1 and 4 Bq m^{-3} was then subtracted from each indoor radon measurement.

make any significant difference of the distribution. The 6-month data corrected for temperature and outdoor radon had an arithmetic mean, a geometric mean and geometric standard deviation of 86 Bq m^{-3}, 19 Bq m^{-3} and 5.44 respectively. This geometric mean is smaller than that based on the data without adjustment for outdoor radon, whereas the geometric standard deviation is larger with adjustment for outdoor radon than without. This is because when the data points less than or equal 4.0 Bq m^{-3} were set to a value of 4.1 Bq m^{-3} and causes some distortion in the lower tail of distribution. It should be noted that the 6-month measurements cover a wider range of radon concentrations (down to very low levels) than do the 3-month measurements.

In both Fig. 3a and b, there appear to be two upward curves, one in the upper part and the other in the lower part of the plot, suggesting a mixture of two normal distributions. A possible explanation is that the 6-month repeated data were from a mixture of national and regional surveys. About 1900 of the 6-month repeat measurements were made during a national survey for which the radon concentrations were generally low, and the remainder were from earlier regional surveys for which the radon concentrations were generally high (Wrixon et al., 1988). Table 3 shows summary statistics for data from these national and regional surveys separately. Not only was the geometric mean of the measured concentrations with temperature and outdoor radon corrections much higher in the regional survey than in the national survey, but the corresponding geometric standard deviation was slightly higher in the regional survey data than in the national survey data (4.3 and 3.3 respectively).

We investigated the 6-month data further by analysing the national and regional survey data separately. Fig. 4 shows normal probability plots of radon concentrations after a logarithmic transformation with temperature corrections (Fig. 4a) and with

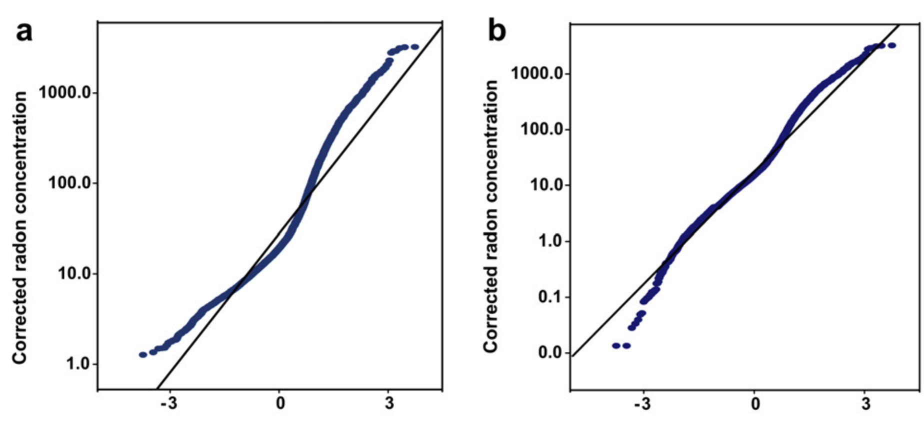

Fig. 3. Normal probability plot of the 6-month repeated radon measurements from 2777 houses, under a logarithmic transformation (a) with temperature corrections and (b) with temperature and outdoor radon corrections.

Table 3
Summary statistics for the corrected 6-month radon concentration data (Bq m^{-3}) from 1908 houses in the national survey and from 869 houses in regional surveys.

	Radon concentrations					
	Min. (Bq m^{-3})	Max. (Bq m^{-3})	Arithmetic mean (Bq m^{-3})	Std. deviation	Geometric mean (Bq m^{-3})	Geometric std. deviation
National survey data						
Temperature corrected	1.27	1295	22.75	48.04	14.73	2.24
Temperature and outdoor radon corrected	0.006	1291	18.80	47.98	9.13	3.30
Regional survey data						
Temperature corrected	2.92	3227	239	337	110	3.75
Temperature and outdoor radon corrected	0.024	3222	235	337	100	4.29

temperature and outdoor radon corrections (Fig. 4b) for the national survey data. In Fig. 4a, there is substantial deviation from linearity in the upper tail of the distribution. When the national data were also corrected for outdoor radon, the points in the upper tail fall closer to a straight line (Fig. 4b), but the lower tail shows departures from the fitted line. This may be because some of the radon measurements were less than the estimated outdoor radon concentration (ie. 4.1 Bq m^{-3}), which might arise because of proportionally large measurement uncertainties at low radon concentrations. Consequently, setting data that were initially less than 4.1 Bq m^{-3} to 4.1 Bq m^{-3} could have led to a deviation in the shape of the lower tail of the plot (Fig. 4b).

Fig. 5 shows the normal probability plot with temperature corrections and with temperature and outdoor radon corrections for the regional survey data. When these data were corrected solely for temperature, the upper and lower tail of the distribution deviates from linearity (Fig. 5a). However, when the data were corrected for outdoor radon as well as for temperature, the data points in lower tail lay closer to the straight line (Fig. 5b), but the pattern in the upper tail does not change. The log-normality in radon measurements and the deviation at the lower and upper tail area are also discussed in details by Bossew (2010).

We also repeated the approach taken in Fig. 2 for the 3-month dataset to see how well repeated 6-month measurements made in the same dwelling agree, both for the national and the regional survey data. Fig. 6 shows a scatter plot of the 6-month data with the regional survey data (Fig. 6a) and the national survey data (Fig. 6b), in both instances with temperature and outdoor radon corrections. We can see that most of the data points from the regional surveys are quite close to the line of equality and that the initial and repeated measurements are highly correlated (correlation coefficient, $r = 0.91$). However, for the national survey data, the fitted regression line does not coincide with the line of equality and a substantial number of data points deviate from the line of equality, especially at low radon concentrations (this point was discussed above). The correlation coefficient for this dataset was 0.66. Fig. 7 shows the plot of the differences between the two logged measurements against the log mean for each dwelling for the regional survey data (Fig. 7a) and for the national survey data (Fig. 7b). Both datasets include temperature and outdoor radon corrections. Both plots show the non-uniform spread of the difference of log radon measurements as a function of averaged log radon measurements and supports the variance of dependent variable varies across the data. For the regional survey data, although there were some data points at low radon concentrations outside the range. The differences against means between the logged measurements were in general distributed around zero (shown by a horizontal grey line in Fig. 7a) i.e. the variability were approximately constant. However, for the national survey data those houses with low radon concentrations displayed a greater variability for these differences, while the variability was constant for those houses with medium to high radon concentrations. Some of the low radon measurements are possible due to relatively high measurement uncertainty at low radon concentration compared with at high concentrations.

2.3. Formal statistical analyses

Repeated measurements of radon concentrations in the same home over a sequence of years have been reported to show

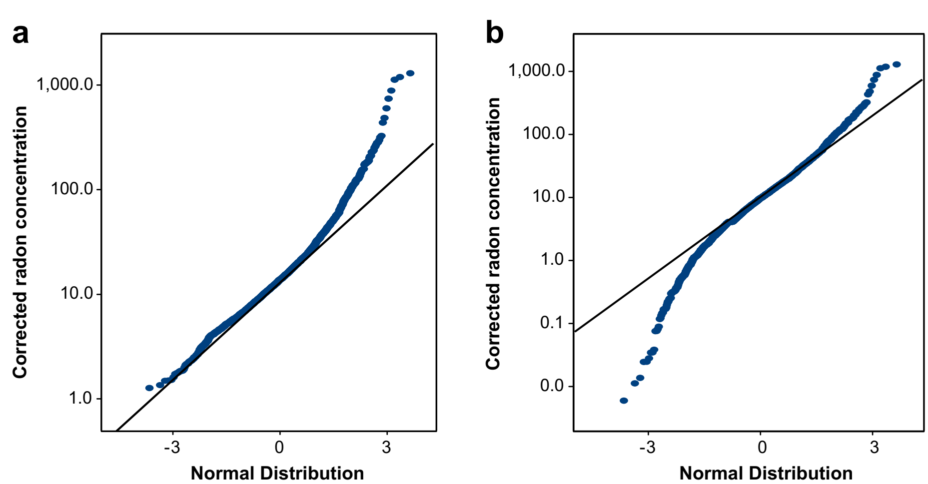

Fig. 4. Normal probability plot of the repeated UK national survey six-month radon measurements from 1908 houses, under a logarithmic transformation (a) temperature corrected, or (b) temperature and outdoor radon corrected.

Please cite this article in press as: Hunter, N., et al., Two error components model for measurement error: application to radon in homes, Journal of Environmental Radioactivity (2011), doi:10.1016/j.jenvrad.2011.05.009

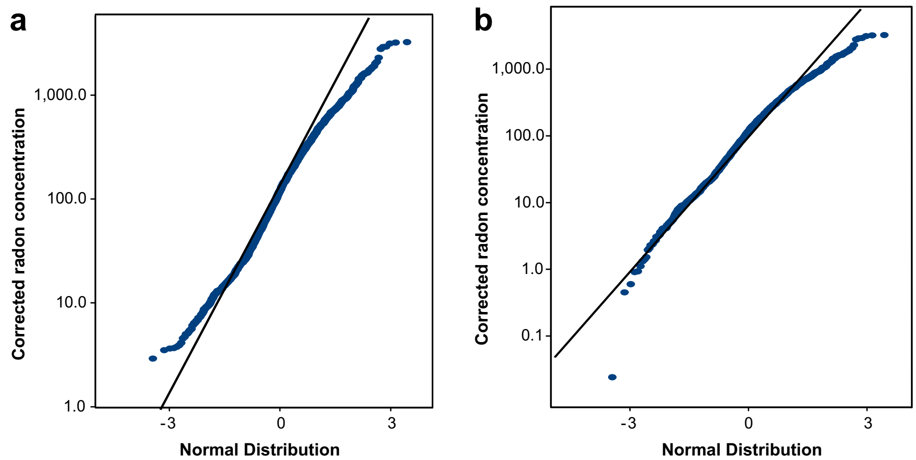

Fig. 5. Normal probability plot of the repeated UK regional survey radon concentrations from 869 houses, under a logarithmic transformation (a) temperature corrected, or (b) temperature and outdoor corrected.

a multiplicative error structure (Darby et al., 2006; Heid et al., 2004, 2006; Hunter et al., 2004; Reeves et al., 1998; Lagarde et al., 1997). Under this multiplicative model, which we will call Model I (Eq. (1)):

$$Y_{ij} = A_i \cdot E_{\text{mul},ij}$$
$$A_i \sim N\left(\theta, \sigma^2_{\text{between}}\right) \text{and } E_{\text{mul}} \sim N\left(0, \sigma^2_{\text{mul}}\right) \quad (1)$$

where Y_{ij} denotes the jth ($j = 1,2$) radon measurement for house i ($i = 1, \ldots, n$; where n is 4566 for the 3-month and 2777 for 6-month repeat measurements), A_i is the underlying radon concentration for house i (distributed as log-normal with mean θ and $\sigma^2_{\text{between}}$) and $E_{\text{mul},\,ij}$ is the random multiplicative error component (distributed as log-normal with mean 0 and σ^2_{mult}). Here θ represents the underlying mean radon level within the study region and σ_{between} represents the between-house standard deviation due to unmeasured or unknown housing structures or other factors.

However, this purely multiplicative error model gives little information on very low concentrations; in reality, as pointed out earlier, there will always be some background noise. The solution for this is a combined model that reflects two types of errors. Under the model proposed in Eq. (2), which we will call Model II (Eq. (2)):

$$Y_{ij} = \left(A_i \cdot E_{\text{mul},ij}\right) + E_{\text{add},ij}$$
$$A_i \sim N\left(\theta, \sigma^2_{\text{between}}\right), E_{\text{mul}} \sim N\left(0, \sigma^2_{\text{mul}}\right) \text{and} E_{\text{add}} \sim N\left(0, \sigma^2_{\text{add}}\right) \quad (2)$$

where $E_{\text{mul},ij}$ represents the proportional error that is exhibited at (relatively) high concentrations and $E_{\text{add},ij}$ (distributed as normal with mean 0 and σ^2_{add}) represents an additive error of background noise that is apparent primarily at very low concentrations. This two error component model incorporates an approximately constant standard deviation at very low concentrations and an approximately constant relative coefficient of variation (CV) at higher concentrations. Furthermore, these two error components are assumed to be independent of one another. Thus, the radon measurement Y is distributed as a linear combination of a normal and a log-normal random variable. Under this model, the variance of Y is:

$$\text{Var}(Y_{ij}) = A^2 S^2_{\text{mul}} + \sigma^2_{\text{add}} \quad \text{where } S_{\text{mul}} = \sqrt{e^{\sigma^2}_{\text{mul}}\left(e^{\sigma^2}_{\text{mul}} - 1\right)} \quad (3)$$

The models in Eq. (1) and Eq. (2) were fitted using the software ‘Bayesian inference Using Gibbs Sampling’ (WinBUGS) (Spiegelhalter et al., 1996). The reason for using this program is that

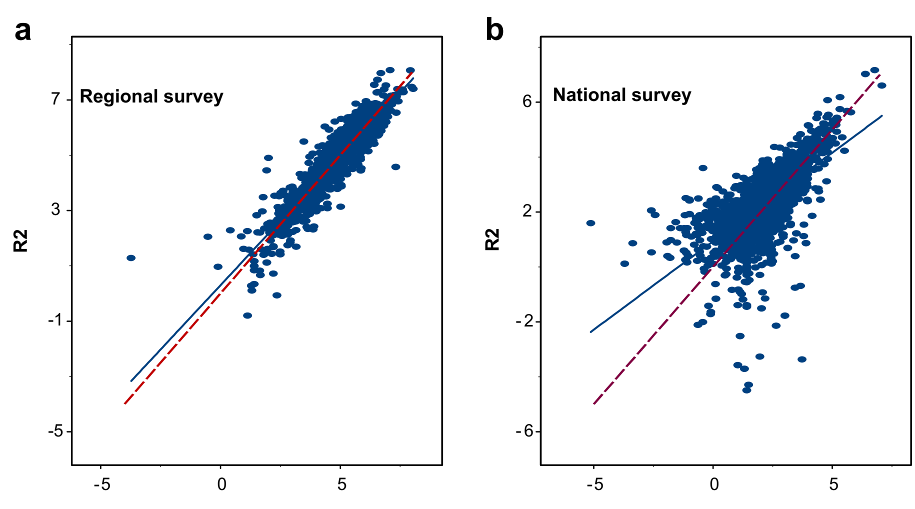

Fig. 6. The relationship between two six-monthly radon measurements, R1 (initial measurement) versus R2 (final measurement); data have been corrected for temperature and outdoor, line of equality (broken line) and a regression line (solid line) for (a) Regional survey data (b) National survey data.

Please cite this article in press as: Hunter, N., et al., Two error components model for measurement error: application to radon in homes, Journal of Environmental Radioactivity (2011), doi:10.1016/j.jenvrad.2011.05.009

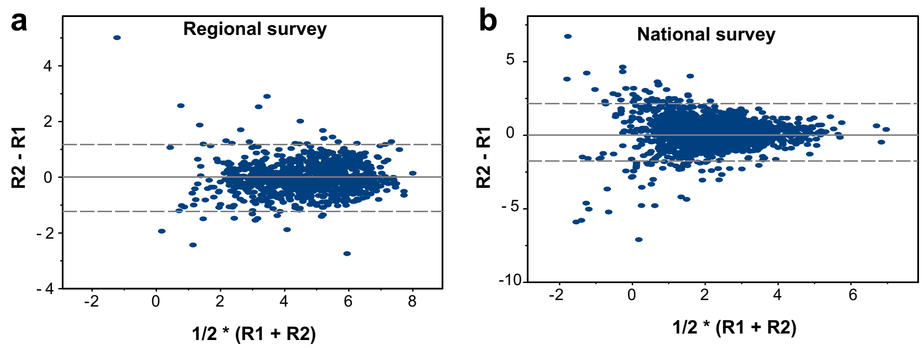

Fig. 7. Difference between the logged two six-monthly radon measurements, plotted against the average of two logged measurements, with the range corresponding to two standard deviations (horizontal dashed lines), for the corrected (a) regional survey data and (b) national survey data.

it provides a powerful means of analysing data using complex models and is available freely (http://www.mrc-bsu.cam.ac.uk/bugs). The program calculates the Deviance Information Criterion (DIC), which can be used to assess model complexity and compare the fit of different models (Spiegelhalter et al., 2002). Hence, the DIC was used in order to select the more appropriate model among those considered. The model with the smallest DIC is estimated to be the model that would best predict a replicate dataset of the same structure as that currently observed. The BUGS code is available from the first author.

3. Results of model fitting

3.1. 3-month repeat measurement data

The estimated values for the overall mean (θ), the between-house standard deviation ($\sigma_{between}$) and the standard deviation of the multiplicative measurement error component (σ_{mult}) using Model I are given in Table 4 for the temperature corrected three month data. The results for the additive (σ_{add}) and multiplicative (σ_{mult}) measurement error components from Model II are also shown in Table 4, together with the DIC values for these two models. It can be seen that allowing for additive measurement error in addition to multiplicative measurement error (Model II) gives a slightly smaller DIC value than allowing for multiplicative measurement error alone.

The estimates of the mean logged radon concentration and of the between-house variance are similar under Models I and II. In contrast, the multiplicative error component is smaller under Model II compared to Model I. The standard deviation of the additive error component is estimated to be $\sigma_{add} = 25$ Bq m^{-3}(95% CI: 23–27). The error component for the between-house variation is smaller than the multiplicative error component under both Model I and Model II. This may be because the initial measurements in the same homes studied were close to the Radon Action Level (200 Bq m^{-3}) and hence, one might expect less variability between homes.

Table 4
Parameter estimates from analysing for 3-month repeated radon data ($n = 4566$) using Model I (Eq. (1)) and Model II (Eq. (2)).

Parameters	[a]Mean estimate (95% CI)	
	Model I	Model II
σ_{mult}	0.41(0.40–0.42)	0.34 (0.33–0.35)
σ_{add} (Bq m^{-3})	–	25 (23–27)
$\sigma_{between}$	0.27 (0.26–0.28)	0.28 (0.26–0.29)
θ	5.13 (5.12–5.14)	5.14 (5.13–5.15)
DIC	105 165	92 507

CI: Confidence interval.
[a] Temperature corrected data.

3.2. Six-month repeat measurement data

The same modelling was applied to the 6-month repeat data. As discussed earlier, these data were split into measurements collected in regional surveys (which generally included high radon values) and those collected in a national survey (which generally contained low radon values).

3.2.1. Regional survey data

Table 5 shows the mean estimates of the error components for the regional survey data with a temperature correction alone and with both temperature and outdoor radon corrections. In both instances, Model II gave a better fit than Model I based on a comparison of DIC values (see Table 5). For the temperature corrected data, the estimates ofσ_{mult}, $\sigma_{between}$ and θ in Model II were very similar to those obtained using Model I (see Table 5). However, for the data with temperature and outdoor radon corrections, the estimated multiplicative error component was slightly smaller under Model II compared to Model I. Under Model II, the estimate of the between-house standard deviation was about four times larger than that for the multiplicative error. This may be because, in general, the repeat measurements for the 6-month regional survey data covered high radon concentrations.

The estimated standard deviation for the additive error component for the regional survey data with temperature correction was slightly smaller than the corresponding estimate from the data with temperature and outdoor radon corrections (see Table 5). This is because correcting outdoor radon in particular equating 4.0 Bq m^{-3} at low concentrations causes some distortion. Hence the additive error component which concentrates mainly at low concentrations expected to be greater for the temperature and outdoor corrected data. However, these estimates ($\sigma_{add} = 1.04$ Bq m^{-3} and $\sigma_{add} = 1.82$ Bq m^{-3}) were not as large as

Table 5
Parameter estimates from analysing 6-month repeated radon data from regional surveys ($n = 869$) using Model I and Model II.

Parameters	Mean estimate (95% CI)			
	[a]Model I	[a]Model II	[b]Model I	[b]Model II
σ_{mult}	0.36 (0.34–0.38)	0.36 (0.34–0.38)	0.42 (0.40–0.44)	0.38 (0.36–0.40)
σ_{add} (Bq m^{-3})	–	1.04 (0.53–1.70)	–	1.82 (1.30–2.46)
$\sigma_{between}$	1.28 (1.21–1.34)	1.27 (1.21–1.34)	1.39 (1.33–1.46)	1.38 (1.31–1.45)
θ	4.70 (4.62–4.79)	4.70 (4.62–4.79)	4.60 (4.51–4.70)	4.61 (4.52–4.71)
DIC	18 585	6969	17 950	6974

[a] Temperature corrected data.
[b] Temperature and outdoor radon corrected data.

Table 6
Parameter estimates from analysing 6-month repeated radon data for national survey data ($n = 1908$) using Model I and Model II.

Parameters	Mean estimate (95% CI)			
	[a]Model I	[a]Model II	[b]Model I	[b]Model II
σ_{mult}	0.39 (0.38–0.40)	0.38 (0.36–0.39)	0.70 (0.68–0.73)	0.50 (0.48–0.52)
σ_{add} (Bq m^{-3})	–	0.85 (0.64–1.05)	–	1.24 (1.08–1.38)
$\sigma_{between}$	0.71 (0.68–0.73)	0.71 (0.68–0.74)	0.96 (0.93–1.00)	0.95 (0.91–0.98)
θ	2.69 (2.65–2.72)	2.69 (2.65–2.72)	2.21 (2.16–2.26)	2.26 (2.21–2.30)
DIC	24 180	13 362	26535	15962

[a] Temperature corrected data.
[b] Temperature and outdoor radon corrected data.

the corresponding estimate for the 3-month repeat measurement data ($\sigma_{add} = 25$ Bq m^{-3}), as given in Table 4. This may be due to the short term 3-month measurements in homes were chosen from among those with larger radon concentrations 200 Bq m^{-3}, while homes in the regional and national survey data cover a wide range values down to very low values.

3.2.2. National survey data

Table 6 shows the mean estimates of the error components for the national survey data with a temperature correction alone and with both temperature and outdoor radon corrections.

Although the estimates of σ_{mult}, $\sigma_{between}$ and θ were similar under Model I and Model II for the temperature corrected data, again Model II gave a better fit than for Model I, as indicated by the smaller DIC value for Model II (see Table 6). The multiplicative error component for the temperature corrected data using Model II, namely $\sigma_{mult} = 0.38$ (95% C.I.: 0.36, 0.39), was similar to the corresponding value for the regional survey data given in Table 5. When the data were corrected for both temperature and outdoor radon, the multiplicative error component using Model II was smaller than that using Model I (see Table 6). For Model II, the estimates of σ_{mult}, σ_{add} and $\sigma_{between}$ based on the temperature corrected data were smaller and the estimate of θ was larger than the corresponding estimates based on the temperature and outdoor radon corrected data (see Table 6). The estimate of the between-house standard deviation under model II was two times larger than that for the multiplicative error based on the national survey data. This was lower than that based on the 6-month regional survey data, possibly because the radon concentrations in the national survey were generally low.

4. Conclusions

Log transformation of the radon data provides good variance stabilisation at high levels, but unrealistically the multiplicative model alone would suggest that the variance at low concentrations was negligibly small. Hence, a simple mixed model that includes two measurement error components, representing additive errors (prominent at low concentrations) and multiplicative errors (prominent at high concentrations), together with a component that represents random variation in radon levels from house to house. The two error (multiplicative and additive) component model (model II here) gave a better fit than a model that includes multiplicative measurement errors alone (Model I) for both the 6-month repeat measurements (from national and regional survey) and the 3-month repeat measurements.

Acknowledgements

The authors would like to thank Chris Howarth (Health Protection Agency) for providing the data.

References

Bossew, P., 2010. Radon: exploring the log-normal mystery. J. Environ. Radioactivity 101, 826–834.
Darby, S., Hill, D., Auvinen, A., et al., 2005. Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. BMJ 330, 223–226.
Darby, S., Hill, D., Deo, H., et al., 2006. Residential radon and lung cancer–detailed results of a collaborative analysis of individual data on 7148 persons with lung cancer and 14,208 persons without lung cancer from 13 epidemiologic studies in Europe. Scand. J. Work Environ. Health 32 (Suppl. 1), 1–83.
Gerken, M., Kreienbrock, L., Wellmann, J., Kreuzer, M., Wichmann, H.E., 2000. Models for retrospective quantification of indoor radon exposure in case-control studies. Health Phys. 78, 268–278.
Gunby, J.A., Darby, S.C., Miles, J.C., Green, B.M.R., Cox, D.R., 1993. Factors affecting indoor radon concentrations in the United Kingdom. Health Phys. 64, 2–12.
Heid, I.M., Kuchenhoff, H., Miles, J., Kreienbrock, L., Wichmann, H.E., 2004. Two dimensions of measurement error: classical and Berkson error in residential radon exposure assessment. J. Expo. Anal. Environ. Epidemiol. 14, 365–377.
Heid, I.M., Kuchenhoff, H., Rosario, A.S., Kreienbrock, L., Wichmann, H.E., 2006. Impact of measurement error in exposures in German radon studies. J. Toxicol. Environ. Health A 69, 701–721.
Hunter, N., Howarth, C.B., Miles, J.C.H., Muirhead, C.R., 2004. Year-to-year variations in radon levels in a sample of UK houses with the same occupants. In: McLaughlin, J.P., Simoloulos, S.E., Stainhäusler, F. (Eds.), Seventh International Symposium on National Radiation in Environment (NRE-VII). Elsevier, Oxford, pp. 438–447.
Hunter, N., Muirhead, C.R., Miles, J.C.H., Appleton, J.D., 2009. Uncertainties in radon related to house-specific factors and proximity to geological boundaries in England. Rad. Prot. Dos 136, 17–22.
Kendall, G.M., Green, B.M.R., Miles, J.C.H., Dixon, D.W., 2005. The development of the UK radon programme. J. Radiol. Prot. 25, 475–492.
Krewski, D., Lubin, J.H., Zielinski, J.M., et al., 2005. Residential radon and risk of lung cancer: a combined analysis of 7 North American case-control studies. Epidemiology 16, 137–145.
Krewski, D., Lubin, J.H., Zielinski, J.M., et al., 2006. A combined analysis of North American case-control studies of residential radon and lung cancer. J. Toxicol. Environ. Health A 69, 533–597.
Lagarde, F., Pershagen, G., Akerblom, G., et al., 1997. Residential radon and lung cancer in Sweden: risk analysis accounting for random error in the exposure assessment. Health Phys. 72, 269–276.
Lubin, J.H., Wang, Z.Y., Boice Jr., J.D., et al., 2004. Risk of lung cancer and residential radon in China: pooled results of two studies. Int. J. Cancer 109, 132–137.
Miles, J.C.H., Appleton, J.D., Rees, D.M., Green, B.M.R., Adlam, K.A.M., Myers, A.H., 2007. Indicative Atlas of Radon in England and Wales. HPA, Didcot. http://www.hpa.org.uk/webw/HPAweb&HPAwebStandard/HPAweb_C/1204186222060?p=1197637096018.
Miles, J.C.H., 1998. Mapping radon-prone areas by log-normal modelling of house radon data. Health Phys. 74, 370–378.
Miles, J.C.H., Howarth, C.B., 2000. Validation Scheme for Laboratories Making Measurements of Radon in Dwellings. NRPB-M1140.
Reeves, G.K., Cox, D.R., Darby, S.C., Whitley, E., 1998. Some aspects of measurement error in explanatory variables for continuous and binary regression model. Stat. Med. 17, 2157–2177.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van der Linde, A., 2002. Bayesian measures of model complexity and fit. J. Royal Stat. Society Ser. B. 64, 1–34.
Spiegelhalter, D.J., Thomas, A., Best, N.G., Gilks, W.R., 1996. BUGS: Bayesian Inference Using Gibbs Sampling, Version 0.50. Technical Report. Medical Research Council Biostatistics Unit, Cambridge.
Wrixon, A.D., Green, B.M.R., Lomas, P.R., et al., 1988. Natural Radiation Exposure in UK Dwellings. NRPB R190. HMSO, London.

Please cite this article in press as: Hunter, N., et al., Two error components model for measurement error: application to radon in homes, Journal of Environmental Radioactivity (2011), doi:10.1016/j.jenvrad.2011.05.009