Lawrence Berkeley National Laboratory

LBL-35819
UC-600

APPARENT NON-LINEARITY IN RESIDENTIAL RADON CONCENTRATION MEASUREMENTS, CAUSED BY THE REGRESSION EFFECT

Phillip N. Price

Energy and Environment Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

December 1994

This work was supported by the Office of Energy Research, the Office of Health and Environmental Research, and the Environmental Sciences Division of the U.S. Department of Energy (DOE) under Contract DE-AC03-76SF00098.

DISCLAIMER

Apparent Non-Linearity in Residential Radon Concentration Measurements, Caused By The Regression Effect: A Comment On Klotz et al. [1]

P. N. Price

Indoor Environment Program 90-3058

Lawrence Berkeley Laboratory

Berkeley, CA 94720

LBL document number 35819

December 6, 1994

Abstract

A recent paper by Klotz et al.[1] examines the relationship between four-day charcoal canister radon measurements and year-long alpha-track detector measurements in 983 New Jersey homes. The ratio of canister measurement to long-term measurement for the homes in the survey, a common parameter of interest, was found to increase as the canister measurement increased. The paper presented considerable discussion of the variation of the ratios as functions of various parameters. Although we did not examine the raw data used in [1], it appears that many or all of the results of Klotz et al. (and perhaps those in other papers [2], [3], [4]), are consistent with a simple model in which both the long-term and short-term measurements provide measurements with error of

the annual-average radon concentration in the home, with no nonlinearity or other functional dependence on radon concentration. We provide an example and discussion of this result, which is caused by the widely known but frequently misunderstood phenomenon called "regression toward the mean", or simply the "regression effect". This in no way invalidates the work in Klotz et al. or in the other papers cited above; we merely wish to bring attention to the fact that the results in these papers may have a very simple explanation.

The typical ratio of short-term to long-term measurement is a common parameter of interest, since short-term measurements are often performed to check whether a home has a radon "problem", where the important parameter for such a question is the long-term radon concentration in the home. The variation of this ratio as a function of one of the measurements is also important, since the appropriate ratio must be chosen if a valid extrapolation from short-term to long-term is to be obtained.

The paper of Klotz et al.[1] presents data and analysis that bear directly on the appropriate ratio for New Jersey homes. The analysis is useful and appropriate, but may have left readers of the article confused as to the causes and implications of the variation of the ratio as a function of canister measurement. In the current work, we attempt to clear up this potential confusion.

We demonstrate below that no matter what tests are performed (even two long-term measurements, for instance), the ratio of the measurement from test A to measurement from test B is essentially guaranteed to vary as a function of test A measurement. The fact that the ratio of measurements varies as a function of one of the measurements does not, therefore, necessarily indicate any non-linearity in the relationship between the tests.

This is perhaps most easily illustrated with an example from the social sciences: imagine a group of students, each given two nearly equivalent tests—perhaps two versions (called A and B) of the famous Scholastic Aptitude Test. We imagine that each student has some "true" test score—the average score that would be obtained if the student took many different versions of the SAT. On any particular version, however, the score will differ from the student's true score: sometimes the student will make a few lucky guesses, or will happen to be asked questions for which she knows the answers; other times the student will be unlucky.

Suppose we examine the group of students who scored extremely high on version A. This group will contain a disproportionate number of "lucky" students, whose test A score is higher than their true score. For most of these students, their test B score will be closer to their true score, and hence lower than their test A score—only a few will be "lucky" twice. So the average ratio of scores (test A / test B) will be greater than unity for this group. Similarly, the ratio of scores for students who did very poorly on test A will be less than unity—a disproportionate number of these students are those who were unlucky (i.e. scored lower than their true score) on test A, and most of these students will do better on test B. This phenomenon is known as "regression towards the mean". The magnitude of the effect depends on the distribution of measured scores (scores on test A and test B) about the true scores: if the test scores can differ greatly from the true scores, the effect can be quite large.

In the above example, whether or not the ratio is of interest depends on the calculation that is being made: if the problem is to estimate the test B score based on the test A score, then the observed ratio should be used. If the problem is to combine data from the tests to obtain a better estimate of the student's true score, then no ratio should be used as a correction factor: the tests are equivalent, and the test scores should simply be averaged. In any case, the fact that the ratio is not constant is not a matter for concern—indeed, it's hard to imagine a situation in which the ratio could be made constant.

In the case of the radon data presented in Klotz et al., the ratios are to be used to predict long-term measurements from short-term measurements, so the examination of these ratios is legitimate. However, it should be realized that most or all of the variation in ratios as a function of canister measurement is probably due to regression towards the mean, and not to some nonlinear relationship between the tests. Indeed, such variation is *inevitable* given almost any realistic model of the measurement process. For example, there is considerable temporal variation in the charcoal-canister measurements. In some homes, the measurement was made over a period for which the indoor concentration is higher than the seasonal average; in other homes, the measurement happened to take place at a low period. A consequence of this variation in measurement is a variation in the ratio of short- to long-term measurement.

We examine a specific model that reproduces some of the statistical features of the analysis of Klotz et al. We do not claim this statistical model is an accurate or adequate description of reality; it is purely illustrative. The model is defined as follows.

1. True annual-average living-area radon concentrations are lognormally distributed, with a geometric mean (GM) of 19 $\mathrm{Bq\,m^{-3}}$ and a geometric standard deviation (GSD) of 2.04. We will find it convenient to work in logarithmic space: let y represent the logarithm of the true radon annual-average living-area radon concentration. Then

$$y \sim \mathrm{N}(\log(19), \log(2.04)), \tag{1}$$

where $\mathrm{N}(a, b)$ represents a normal distribution with mean a and standard deviation b.

2. The living-area alpha-track measurements are lognormally distributed about the true values with a geometric standard deviation of 1.5 (i.e., measured values typically differ from the true value by about 50% or less). With the logarithm of the measurement denoted by α, we have

$$\alpha = y + \mathrm{N}(0,\log(1.50)). \tag{2}$$

In untransformed space, this means the measured value is equal to the true concentration multiplied by a lognormally distributed error term.

3. The basement canister measurements are biased by a factor of 3.1 from the true annual-average living-area concentrations, and vary lognormally about this biased value with a geometric standard deviation of 1.8. With the logarithm of the basement canister measurement denoted by β, we have

$$\beta = \log(3.1) + y + \mathrm{N}(0,\log(1.8)). \tag{3}$$

In untransformed space, this implies that the basement canister measurement is equal to 3.1 times the annual living-area concentration multiplied by a lognormally distributed error term. This error term includes all variation between measurement and true concentration, such as temporal variation, house-to-house variation in the true ratio, and measurement error.

The model described above is the simplest plausible model: true radon concentrations are lognormally distributed, and both measurements are proportional to the true concentration but with a lognormally distributed fractional error. We emphasise that we do not claim that this model is an accurate reflection of reality; specifically, one could quarrel with the selection of a geometric standard deviation of 1.5 for the alpha-track detectors errors, since quality-assurance data mentioned in Klotz et al. (based on collocated detectors) suggest that the typical error should vary from 1.25 to 1.1 or lower as the alpha-track measurement increases.

The parameters in the current model were chosen merely to provide agreement with the relevant parts of Tables 1 and 3 in Klotz et al. The key qualitative aspect of the analysis—the increasing ratio as a function of basement measurement—will occur as long as there is any independent variability between the basement measurement and the alpha-track measurement, whether due to instrumental error, temporal variation, or whatever.

A simulation of 4000 sets of "measurements" based on the model yielded the results in Table 1, which shows the geometric mean and geometric standard deviation of the ratio of basement measurement to living-area measurement for different concentration bins; the corresponding results from Table 3 in Klotz et al., are shown as well. The overall GM and GSD of the simulated alpha-track measurements were 19.1 $Bq.m^{-3}$ and 2.3, respectively; for the simulated canister measurements, they were 59 $Bq.m^{-3}$ and 2.5. All of these figures are approximately in agreement with the data in Klotz et al.

As the table shows, even the simple, linear model described above reproduces the variation of ratio GMs extremely well. This does not alter the primary conclusions of the work of Klotz et al.; for example, basement measurements greater than 150 $Bq.m^{-3}$ are likely to be about 5.6 times higher than the living-area concentration. But we do think an understanding of the cause of the variation in the ratio is necessary if the results are to be used for substantive predictions. For instance, an attempt to analyze different datasets, based on measurements for which the variance parameters are different, will be meaningful only if a correct statistical model is used.

Basement measurement (Bq/m^3)	ratio, from Klotz et al. GM (GSD)	ratio, from current model GM (GSD)
< 37	1.8 (2.0)	2.0 (1.9)
37–73	2.9 (1.8)	2.9 (1.9)
74–149	4.6 (2.0)	3.8 (1.9)
150–299	5.5 (1.9)	5.3 (1.9)
300+	6.5 (2.8)	6.8 (1.8)
< 150	2.7 (2.2)	2.8 (2.0)
150+	5.6 (2.0)	5.7 (1.9)
Total	3.1 (2.3)	3.1 (2.1)

Table 1: Ratio of short-term basement to long-term living level radon concentrations, for various basement measurements. Results from Klotz et al. and from the current model are shown.

We have not attempted to create a model to reproduce all of the ratios given in the other tables of Klotz et al., but it is apparent that the effects described in the current work can produce behavior of the correct magnitude.

On a related subject, we note that although the model assumes a simple multiplicative relationship between *true* basement and living-level concentrations, a linear regression of the logarithm of living-level *measurement* on the logarithm of the basement *measurement* will yield a slope of less than one, corresponding to a nonlinear relationship between the two measurements. (With the current model the slope is 0.74). Such a nonlinearity has been previously noted in radon measurements, as in White et al. [3].

We agree with the conclusion of Klotz et al. that the variation in the ratio has important implications in attempting to estimate the true distribution of radon exposures based on screening measurements. Basically, extremely high measurements are likely to be too high, and extremely low measurements are likely to be too low, compared to the true annual-average concentration in the home. A better estimate for the true annual-average concentration is provided by bringing all of the values towards the center of the distribution, by an amount that depends on the magnitude of the measurement; the ratios determined by Klotz et al. are empirical estimates of the amount by which this "drawing in" of the distribution should be done. We note that an existing technique, known as Bayesian Hierarchical Modeling (see, for example, [5], [6], [7]), is very well-suited to the analyses that are required for this sort of work, and is on a somewhat sounder and better-understood statistical footing.

References

[1] Klotz, J. B.; Schoenberg, J. B.; Wilcox, H. B. Relationship Among Short– and Long–Term Radon Measurements Within Dwellings: Influence of Radon Concentrations. Health Phys. 65(4): 367–374; 1993.

[2] Perritt, R.L.; Hartwell, T.K.; Sheldon, L.S.; Cox, B.G.; Clayton, C.A.; Jones, S.M.; Smith, M.L.; Rizzuto, J.E. Radon-222 levels in New York state homes. Health Phys. 58:147–155; 1990.

[3] White, S.B.; Clayton, G.A. ; Alexander, B.V.; Clifford, M.A. A statistical analysis: Predicting annual Rn-222 concentrations from 2-day screening tests. In: Proceedings of the 1990 Symposium on Radon and Radon Reduction Technology. Washington, D.C.: U.S. EPA; Vol 1; 1990: 3-117–3-118.

[4] Ronca-Battista, M; Chiles, B. The relationship between winter screening and annual average radon concentrations in U.S. homes. In: Proceedings of the 1990 International Symposium on Radon and Radon Reduction Technology. Washington, D.C.: U.S. EPA; Vol 1; 1990a:2-3 – 2-14.

[5] Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis.* Wiley Classics, Wiley & Sons, New York.

[6] Rubin, D. B. (1980). Using empirical Bayes techniques in the law school validity studies (with discussion). *Journal of the American Statistical Association* **75**, 801–827.

[7] Gelman, A; Carlin, J; Stern, H; Rubin, D. (1994) *Bayesian Data Analysis*, to be released as a textbook by Chapman and Hall, New York.