

Interactive comment on “Post Chernobyl surveys of radiocaesium in soil, vegetation, wildlife and fungi in Great Britain” by J. S. Chaplow et al.

Anonymous Referee #2

Received and published: 28 March 2015

Review Chaplow et al., ESSD-2014-27

I applaud the decision by the authors and CEH to make this data publicly available. I thank the authors for their efforts. I understand fully that these data represent the post-hoc outcome of sampling strategies driven by immediate and evolving political and scientific concerns, never intended to represent systematic long-term monitoring. I also expect that after the initial outcry and interest, the authors will have encountered substantial funding and logistical barriers that limited what they might have wished to accomplish.

Despite the historical and present day value of this data, and expected interest from many directions, the data sets and their descriptions require substantial modification

C361

and improvement to improve access and explanation and to demonstrate reliability.

First, we must have these data geo-referenced in ISO lat lon coordinates, not in northings and eastings (for which the authors specify no reference baseline!). Incidentally, in data set 4, the authors have the wrong column headings (columns C, D and E).

(One wonders, unfortunately, about the accuracy and consistency of other column headings. Please can the authors check for other mistakes?)

Likewise, we need ISO year month day. Data set 5 includes this information but not in a consistent or ISO format.

The codes, column A in each data set, provide distraction and, apparently, no useful or consistent information. They evidently tie, albeit without consistency, to local place names. Delete these or replace them with some useful designations. Despite what the authors say (Page 5, line 4) about supporting metadata, I could find no useful metadata to help explain the .csv files either on the CEH landing page or on the doi link page. Useful metadata needs to occur in the document, not in some external and uncertain destination.

A few soil data appear to emerge as the only common factor among all the five data sets. But because each data set has its own counting techniques and uncertainties, we have no way to know whether we can consider those data comparable across data sets. Please put the counting statistics, together with the handling (sampling, drying, weighing) uncertainties together, all in one section. Of course detectors improve, software changes, standards evolve, even for a time apparently they used some faster but lower-precision gamma detectors to pre-screen samples and thereby focus the longer-time measurements on the higher concentration samples.

We see data for several other isotopes, of Ru, K, Pu, etc. All gamma emitters I assume (I didn't look them up) and all sampled and counted by similar techniques? If these extra data have value, then they deserve a discussion of measurement protocol. I

C362

would tend to keep them in but only with adequate description. Otherwise, leave them out.

The reference list seems a little odd, covering peer-reviewed literature, technical reports and newspaper articles. One hopes that it offer a unique historical view?

Specific comments follow.

Page 2, line 6 "Additional data to radiocaesium" should read 'additional data for radio-caesium'?

Page 2, line 11 - Downloading data. I did successfully penetrate the CEH site - as it turns out I already had a CEH account - to download the set of .csv files. At no cost. But this does not represent open access as ESSD advertises? Too many steps and barriers. And now CEH knows my name or at least my email address, which potentially compromises the anonymity of this review. If the UK mandates these access barriers for their publicly funded data centres (I notice similar barriers at BADC) then the authors really should consider putting the data elsewhere, e.g. Pangaea. This reviewer wonders whether ESSD should enforce some data access standards?

Page 6, lines 10 to 14 - With the mixture of common and species names and in several cases more than one species, this section reads a bit confusing especially if one does not recognize the species names. A slightly revised format would make the lists clearer, e.g.:

other vegetation types - heather (*Calluna vulgaris*) and bracken (*Pteridium aquilinum*) (Horrill et al., 1988). In addition to vegetation, samples of wildlife - rabbit (*Oryctolagus cuniculus*), hare (*Lepus capensis* and *Lepus timidus*), fox (*Vulpes vulpes*), red deer (*Cervus elaphus*), woodcock (*Scolopex rusticola*) and grouse (*Tetrao tetrix* and *Lagopus lagopus*) - were collected.

From table A, 1585 samples, confirmed in data set 1.

Data set 1

C363

May 1986: 16 land classes, 20 squares within each land class, 320, eliminate 2, 318 independent samples. But data set 1 .csv file contains 668 data rows all designated as grass but with N, C, X and Z codes. Even those with obvious dry weights (e.g. indicative of biomass samples) include 287 designated as C and 301 designated as N. Please explain?

Repeated in October 1986, and extended to include animals: At face value 318 vegetation plus 146 animal tissue plus 94 gut contents gives 558 independent samples. .csv file contains 615 data rows, variously C, N, X, Y, Z.

Spring 1987, reanalyze 100 stations in Scotland plus 56 additional. .csv file contains 302 data rows, now variously A, B, C, D and N.

$668 + 615 + 302 = 1585!!$ Where do all the 'extra' samples come from. A geographical and graphical time line, showing what the authors measured when and where - not as in Figure 2! - would prove very helpful.

Page 6, line 26 - so dry weights in column F represent dry biomass per meter squared, not sampled biomass (e.g. 40 to 50 g)?.

Page 7, lines 15 to 20 - Please we need some information about precision and accuracy. Detectors efficiency as low but respectable for relatively long counting times. Did the authors run these samples to some minimal counting time to achieve minimum counts or to achieve reliable counting statistics.

E.g for May 1986 samples, all designated N, range 1 to nearly 16,000 Bq, some with .1 apparent precision, what can we say about certainty or uncertainty of individual data point?

Page 7, line 22 - corrected to date of sampling. This implies that authors know exact sampling date for each sample. Include this in the data?

Rows 1369 to 1391 - what sample type? Ditto for rows 1443 to 1486 - what sample type?

C364

Figures 1 and 2 - label needs correction, don't understand the copyright - applies to the maps themselves? Can one have copyrighted content within an open access product? I doubt it. In any case, we as reviewers and users should have the ability to reproduce these figures from the data provided (if we had usable geo-reference data, see my comment above) and in fact the authors should have done likewise, not simply repeat an earlier figure. Refresh this and remove the copyright.

Page 8, lines 2,3 - Data set 2, samples redundant with spring 1987 of data set 1?

Page 8, lines 4,5 - How did the additional detectors change or improve the counting statistics?

Page 8, line 25 - Cesium from other sources? Need the evidence? High ratios? Low ratios? Citation for this?

Page 8, line 26 - Sigmoidal relationship rain to Cs137. What kind of rainfall data? We don't have access to that data here? Why include this? Citation for this?

Page 9, lines 19 to 23 - Here we first get counting statistics. Do these apply to earlier data as well? Should we have an overall error / uncertainty section? Awkward phrasing and punctuation. Citation for this error analysis?

Page 9, line 24 - The standards used here differ from those used earlier?

Page 10, line 2 - similar to what? Similar for Chernobyl sources vs test sources?

Page 10, line 8 - farm codes A, B and C. But C used extensively as a code for prior samples?

Page 10, lines 25-27 - here we find different counting statistics, apparently using more efficient detectors.

Page 11, lines 1,2 - this represents a different counting error statistic than specified before?

C365

Page 12, lines 5,6 - yet another counting criteria, including "where possible".

Data set 2 - No dates! Different label codes? Rapid counting to eliminate samples below 300, 340? What differences for rows 338 to 382?

Data set 3 - Approximate dates. More isotopes including plutonium? Why included? Description implies vegetation biomass data but not included or not distinguished (e.g. no biomass numbers)?

Data set 4 - Labels for columns C, D and E are wrong? Other columns as well? Column G, vegetation group, similar to land classes used earlier? Vegetation biomass much smaller, by several orders of magnitude, to data in data set 1? Strange characters in column F?

Data set 5 - day month year labels but inconsistent format. Fungi data by species, but also by kg dry weight? Soil data included but very scattered?

Interactive comment on Earth Syst. Sci. Data Discuss., 7, 693, 2014.

C366