Case Study 2

Parker Johnson

11/8/2022

Introduction

Radon levels have been a subject of concern for homeowners in Minnesota, as too high of radon levels can become dangerous and even lead to lung cancer. Therefore, having an understanding of radon levels in Minnesota and how they might differ within each county is important for everyone's well-being. In this study, we estimate county-level average radon concentrations, as well as the state-wide average radon concentration. From this, we then identify counties with unusually high radon concentrations.

Methodology

For this estimation, a hierarchical model was used. This model seemed appropriate for the study, because a hierarchical model can acknowledge the grouping structure of the counties while still addressing the broader population of Minnesota at the same time.

I will assume that the log radon levels are normally distributed, since log levels can be both positive and negative. In this situation, Y_{ij} represents the i-th radon level for county j. This hierarchical model will need μ_j, which is the mean of the log radon levels for county j. I will assume a common standard deviation across counties, since the spread of log radon values across the state doesn't seem likely to differ much across counties, and this allows for a simpler implementation. I will also work with precision, which is $\frac{1}{\sigma^2}$, within this model to allow for easier implementation before transferring this back to standard deviations. Precisions often have gamma priors; therefore, the hierarchical model will have the following elements:

$$\text{Sampling model: For j = 1, ... 85 and i = 1, ..., } n_j : Y_{ij}|\mu_j, \sigma \overset{\text{iid}}{\sim} Normal(\mu_j, \sigma)$$

$$\text{Prior for } \mu_j\text{, Stage 1: j = 1, ..., 85: } \mu_j|\mu, \tau \sim Normal(\mu, \tau)$$

$$\text{Prior for } \mu_j\text{, Stage 2: } \mu, \tau \sim \pi(\mu, \tau)$$

$$1/\sigma^2|a_\sigma, b_\sigma \sim Gamma(a_\sigma, b_\sigma)$$

Where μ and τ are random hyperparameters and $\pi(\mu, \tau)$ denotes an arbitrary joint hyperprior distribution.

To construct a prior on these hyperparameters, a typical approach for Normal models is to assign two independent prior distributions; a Normal distribution for the mean μ and a Gamma distribution for the precision $1/\tau^2$. Therefore, we now have the following prior for μ_j instead of $\pi(\mu, \tau)$:

$$\mu|\mu_0, \gamma_0 \sim Normal(\mu_0, \gamma_0)$$

$$1/\tau^2|a, b \sim Normal(a_\tau, b_\tau)$$

We now need to choose values for $a_\sigma, b_\sigma, \mu_0, \gamma_0, a_\tau$, and b_τ. According to https://www.epa.gov/radon/what-epas-action-level-radon-and-what-does-it-mean#:~:text=The%20average%20indoor%20radon%20concentration,related%20lung%20cancers%20a%20year, the average indoor radon concentration for homes in the US is about 1.3 pCi/L. On the log scale, this value is approximately 0.26, so this will be the value for μ_0. Since I don't know how variable this value is, I will choose $\gamma_0 = 1$ to keep a fair amount of variability. Weakly informative priors are sensible to choose for τ and σ, so I will choose $a_\sigma = 1, b_\sigma = 1, a_\tau = 1$, and $b_\tau = 1$. Therefore, the complete hierarchical model is as follows:

Sampling model: For j = 1, … 85 and i = 1, …, n_j :

$$Y_{ij}|\mu_j, \sigma_j \overset{\text{iid}}{\sim} Normal(\mu_j, \sigma_j)$$

Prior for μ_j, Stage 1: j = 1, …, 85:

$$\mu_j|\mu, \tau \sim Normal(\mu, \tau)$$

Prior for μ_j, Stage 2: the hyperpriors:

$$\mu \sim Normal(0.26, 1)$$

$$1/\tau^2 \sim Gamma(1, 1)$$

Prior for σ:

$$1/\sigma^2 \sim Gamma(1, 1)$$

With the full model, JAGS can now be used to estimate the posterior distribution in order to estimate county-level and state-wide radon concentrations. The counties with unusually high radon concentrations were found by identifying the top 3 counties for their mean radon values, and by also identifying the top 3 counties for the 97.5th quantile for their radon values. The 97.5th quantile counties were identified on top of the mean radon values because this means that it's still possible for the mean log radon value to be this high in these given counties, so the uncertainty in these counties for a high mean value means that a threat for high radon values could still be present.

Results

The posterior mean of the log radon level in each county from JAGS can be seen in Figure 1. From these posterior means, the largest estimated mean was in Blue Earth County, with a mean of 1.777 log pCi/L (which is equivalent to 5.909 pCi/L). The smallest estimated mean was in Lake County, with a mean of 0.643 log pCi/L (which is equivalent to 1.903 pCi/L).

The posterior mean of the log radon level across every county is approximately 1.317 log pCi/L (which is equivalent to 3.732 pCi/L), with a standard deviation of approximately 0.0557 log pCi/L. Therefore, there is a 95% chance that the state-wide average radon level is between 3.354 and 4.163 pCi/L.

The counties with the 3 highest mean radon values were Blue Earth, Lac Qui Parle, and Freeborn County, with respective means of 1.777 log pCi/L (or 5.909 pCi/L), 1.741 log pCi/L (or 5.703 pCi/L), and 1.740 log pCi/L (or 5.697 pCi/L). The counties with the 3 highest radon values for their 97.5th quantiles were Lac Qui Parle, Watonwan, and Nicollet, with respective quantile values of 2.416 log pCi/L (or 11.21 pCi/L), 2.317 log pCi/L (or 10.14 pCi/L), and 2.291 log pCi/L (or 0.888 pCi/L). It's worth noting that these 3 counties all had means greater than 1.70 log pCi/L as well.

Discussion

The state-wide mean log radon concentration is approximately 1.317 log pCi/L (which is equivalent to 3.732 pCi/L), and a table of the average mean log radon concentrations in each county can be seen after the discussion section. From these results, we estimate that the counties with unusually high radon values include Blue Earth, Lac Qui Parle, Freeborn, Watonwan, and Nicollet.

According to https://www.consumerreports.org/radon/is-it-safe-to-buy-a-home-with-an-elevated-radon-level/, levels of 4 pCi/L, or 1.386 log pCi/L, or higher are considered hazardous. However, radon levels below this level can still pose a threat to the health of people living in that home. Since the state-wide mean radon concentration is fairly close to this level, efforts should be taken across the state as a whole to reduce radon levels in homes for the general health of Minnesotans. People living in the counties of Blue Earth, Lac Qui Parle, Freeborn, Watonwan, and Nicollet should be particularly cautious about radon gas, and should make sure that their home's radon levels are not hazardous.

One drawback to this design is the possibility for inaccuracies to arise when justifying the prior distribution and model. While a weakly informative prior does not have a large impact on the posterior distribution, it does sway it slightly, so an incorrect prior distribution will result in a sightly inaccurate posterior distribution

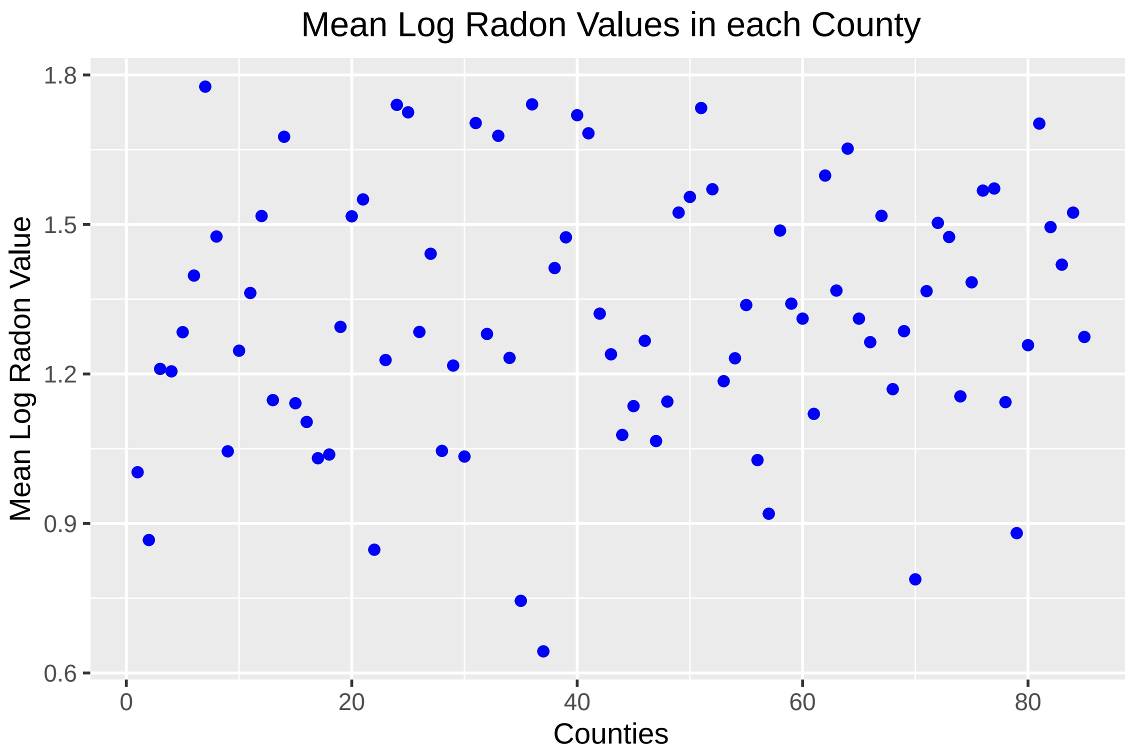

Figure 1: Scatterplot of mean log radon values in each county.

and interpretations. Since the country-wide radon mean level might differ heavily from Minnesota's mean level, perhaps more research could have been conducted to find a more accurate prior distribution for Minnesota.

Another drawback to this design is that there isn't a whole lot of observed data for many of the counties. Mahnomen, Murray, and Wilkin only had 1 observed radon level, with Murray and Wilkin's value being very high. However, with only 1 observation, a good estimation of the mean radon level in those counties can't be obtained, so it's hard to know whether the radon levels in these counties is hazardous or not. Additionally, many other counties had less than 10 observations, so these lack of observations adds uncertainty to the posterior means that are obtained and the strength of the county-wide means are weakened. Therefore, if possible, redoing this study with more observations across each county would be beneficial to truly understanding radon levels within each county and across Minnesota.

Table 1: Table of mean log radon values in each county and their quantiles.

	county	Mean	SD	2.5th quantile	25th quantile	50th quantile	75th quantile	97.5th quantile
mu_j[1]	AITKIN	1.003	0.288	0.434	0.811	1.004	1.197	1.569
mu_j[2]	ANOKA	0.867	0.105	0.663	0.796	0.867	0.937	1.072
mu_j[3]	BECKER	1.210	0.299	0.620	1.006	1.212	1.417	1.777
mu_j[4]	BELTRAMI	1.205	0.239	0.743	1.044	1.207	1.365	1.677
mu_j[5]	BENTON	1.284	0.280	0.737	1.096	1.284	1.469	1.825
mu_j[6]	BIG STONE	1.397	0.298	0.823	1.195	1.395	1.599	1.983
mu_j[7]	BLUE EARTH	1.777	0.187	1.404	1.650	1.776	1.899	2.152
mu_j[8]	BROWN	1.476	0.283	0.933	1.284	1.471	1.669	2.026
mu_j[9]	CARLTON	1.045	0.213	0.626	0.900	1.046	1.189	1.463
mu_j[10]	CARVER	1.247	0.250	0.761	1.078	1.249	1.414	1.733

	county	Mean	SD	2.5th quantile	25th quantile	50th quantile	75th quantile	97.5th quantile
mu_j[11]	CASS	1.362	0.263	0.853	1.188	1.360	1.539	1.884
mu_j[12]	CHIPPEWA	1.517	0.281	0.957	1.333	1.520	1.705	2.066
mu_j[13]	CHISAGO	1.148	0.251	0.659	0.982	1.146	1.316	1.635
mu_j[14]	CLAY	1.676	0.185	1.304	1.553	1.676	1.801	2.041
mu_j[15]	CLEARWATER	1.141	0.279	0.584	0.947	1.141	1.329	1.695
mu_j[16]	COOK	1.104	0.325	0.465	0.888	1.113	1.321	1.736
mu_j[17]	COTTONWOOD	1.031	0.283	0.474	0.842	1.032	1.224	1.574
mu_j[18]	CROW WING	1.038	0.196	0.653	0.909	1.038	1.173	1.418
mu_j[19]	DAKOTA	1.294	0.099	1.102	1.228	1.293	1.363	1.488
mu_j[20]	DODGE	1.516	0.302	0.943	1.308	1.513	1.722	2.132
mu_j[21]	DOUGLAS	1.550	0.218	1.123	1.401	1.549	1.694	1.987
mu_j[22]	FARIBAULT	0.847	0.250	0.357	0.684	0.849	1.018	1.324
mu_j[23]	FILLMORE	1.228	0.324	0.586	1.017	1.226	1.446	1.861
mu_j[24]	FREEBORN	1.740	0.221	1.315	1.588	1.741	1.887	2.177
mu_j[25]	GOODHUE	1.725	0.189	1.362	1.595	1.724	1.854	2.090
mu_j[26]	HENNEPIN	1.284	0.076	1.140	1.233	1.285	1.335	1.432
mu_j[27]	HOUSTON	1.441	0.249	0.961	1.276	1.435	1.613	1.923
mu_j[28]	HUBBARD	1.046	0.263	0.511	0.874	1.046	1.227	1.550
mu_j[29]	ISANTI	1.217	0.303	0.622	1.015	1.215	1.420	1.816
mu_j[30]	ITASCA	1.034	0.208	0.621	0.895	1.037	1.176	1.443
mu_j[31]	JACKSON	1.704	0.258	1.206	1.524	1.710	1.876	2.207
mu_j[32]	KANABEC	1.280	0.274	0.757	1.094	1.279	1.461	1.817
mu_j[33]	KANDIYOHI	1.678	0.281	1.128	1.489	1.679	1.866	2.228
mu_j[34]	KITTSON	1.232	0.296	0.649	1.033	1.233	1.430	1.818
mu_j[35]	KOOCHICHING	0.745	0.244	0.253	0.581	0.747	0.913	1.211
mu_j[36]	LAC QUI PARLE	1.741	0.339	1.089	1.508	1.741	1.972	2.416
mu_j[37]	LAKE	0.643	0.227	0.198	0.488	0.647	0.796	1.077
mu_j[38]	LAKE OF THE WOODS	1.413	0.274	0.877	1.230	1.412	1.597	1.946
mu_j[39]	LE SUEUR	1.474	0.263	0.946	1.295	1.477	1.649	1.991
mu_j[40]	LINCOLN	1.719	0.284	1.164	1.531	1.714	1.907	2.284
mu_j[41]	LYON	1.683	0.231	1.241	1.519	1.678	1.843	2.127
mu_j[42]	MAHNOMEN	1.321	0.357	0.621	1.087	1.324	1.557	2.031
mu_j[43]	MARSHALL	1.239	0.221	0.804	1.093	1.237	1.388	1.663
mu_j[44]	MARTIN	1.078	0.237	0.617	0.918	1.077	1.237	1.546
mu_j[45]	MCLEOD	1.135	0.190	0.767	1.004	1.135	1.266	1.506
mu_j[46]	MEEKER	1.267	0.269	0.739	1.085	1.261	1.444	1.802
mu_j[47]	MILLE LACS	1.065	0.332	0.410	0.843	1.062	1.285	1.730
mu_j[48]	MORRISON	1.145	0.216	0.714	1.001	1.146	1.290	1.562
mu_j[49]	MOWER	1.524	0.191	1.147	1.395	1.524	1.653	1.899
mu_j[50]	MURRAY	1.555	0.361	0.856	1.316	1.546	1.796	2.277
mu_j[51]	NICOLLET	1.734	0.284	1.180	1.542	1.729	1.926	2.291
mu_j[52]	NOBLES	1.571	0.304	0.991	1.365	1.563	1.771	2.169
mu_j[53]	NORMAN	1.186	0.300	0.595	0.985	1.189	1.383	1.768
mu_j[54]	OLMSTED	1.231	0.151	0.936	1.130	1.231	1.334	1.526
mu_j[55]	OTTER TAIL	1.338	0.231	0.883	1.184	1.337	1.494	1.792

	county	Mean	SD	2.5th quantile	25th quantile	50th quantile	75th quantile	97.5th quantile
mu_j[56]	PENNINGTON	1.027	0.311	0.408	0.828	1.029	1.239	1.636
mu_j[57]	PINE	0.920	0.256	0.428	0.747	0.924	1.094	1.415
mu_j[58]	PIPESTONE	1.488	0.277	0.944	1.303	1.487	1.668	2.040
mu_j[59]	POLK	1.341	0.280	0.790	1.152	1.348	1.530	1.886
mu_j[60]	POPE	1.311	0.317	0.682	1.096	1.313	1.521	1.922
mu_j[61]	RAMSEY	1.120	0.132	0.860	1.030	1.119	1.211	1.379
mu_j[62]	REDWOOD	1.598	0.267	1.086	1.418	1.598	1.773	2.128
mu_j[63]	RENVILLE	1.367	0.302	0.773	1.167	1.366	1.566	1.979
mu_j[64]	RICE	1.652	0.203	1.248	1.518	1.655	1.786	2.050
mu_j[65]	ROCK	1.311	0.325	0.675	1.097	1.310	1.527	1.955
mu_j[66]	ROSEAU	1.264	0.188	0.886	1.139	1.262	1.391	1.629
mu_j[67]	SCOTT	1.517	0.194	1.142	1.388	1.518	1.648	1.900
mu_j[68]	SHERBURNE	1.170	0.228	0.727	1.014	1.168	1.323	1.609
mu_j[69]	SIBLEY	1.286	0.284	0.721	1.098	1.292	1.474	1.843
mu_j[70]	ST LOUIS	0.788	0.073	0.646	0.738	0.787	0.837	0.932
mu_j[71]	STEARNS	1.366	0.144	1.081	1.270	1.366	1.463	1.650
mu_j[72]	STEELE	1.503	0.214	1.077	1.361	1.505	1.642	1.932
mu_j[73]	STEVENS	1.475	0.326	0.852	1.253	1.463	1.693	2.127
mu_j[74]	SWIFT	1.155	0.280	0.603	0.966	1.159	1.341	1.708
mu_j[75]	TODD	1.384	0.300	0.785	1.187	1.386	1.582	1.958
mu_j[76]	TRAVERSE	1.568	0.284	1.018	1.372	1.568	1.760	2.132
mu_j[77]	WABASHA	1.572	0.239	1.109	1.413	1.572	1.731	2.044
mu_j[78]	WADENA	1.143	0.262	0.638	0.966	1.143	1.317	1.669
mu_j[79]	WASECA	0.881	0.285	0.319	0.693	0.887	1.073	1.423
mu_j[80]	WASHINGTON	1.258	0.113	1.039	1.183	1.257	1.332	1.484
mu_j[81]	WATONWAN	1.702	0.309	1.105	1.491	1.702	1.912	2.317
mu_j[82]	WILKIN	1.495	0.358	0.804	1.249	1.489	1.735	2.203
mu_j[83]	WINONA	1.419	0.193	1.044	1.285	1.420	1.551	1.801
mu_j[84]	WRIGHT	1.524	0.194	1.142	1.389	1.523	1.655	1.910
mu_j[85]	YELLOW MEDICINE	1.274	0.320	0.642	1.060	1.276	1.490	1.894