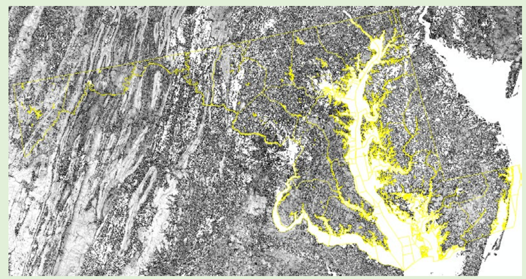
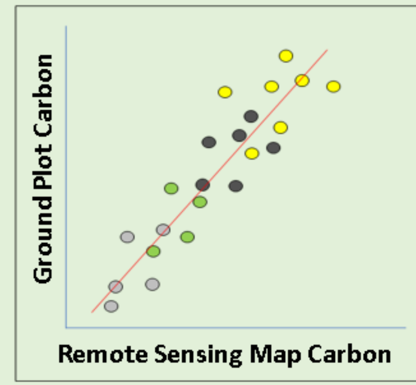
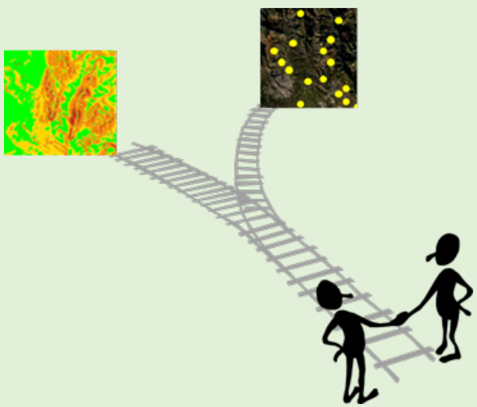
Forest carbon accounting with model-based estimation from remote sensing data and Nationwide Forest Inventory plots

Forest Inventory and Analysis
We are the Nation's Forest Census
Welcome!







Andrew Lister
USDA Forest Service, Forest Inventory and Analysis (FIA)

With thanks to:
Laura Duncanson, UMD
John Hogland, USDA Forest Service
Neha Hunk, UMD
George Hurtt, UMD
Lei Ma, UMD
Barry T. Wilson, USDA Forest Service

What does FIA do?

Reporting on the status of and trends in the Nation's forest resources; technology transfer to NFS and partners; basic and applied research; information for policy….

<u>Land variables</u>
- Forest type
- Site class
- Stand size
- Physiographic class
- Land use
- Land cover
- Owner type/class

…

<u>Tree variables</u>
- Species
- Diameter
- Height
- Merchantability
- Damage type
- Crown ratio
- Growth

…

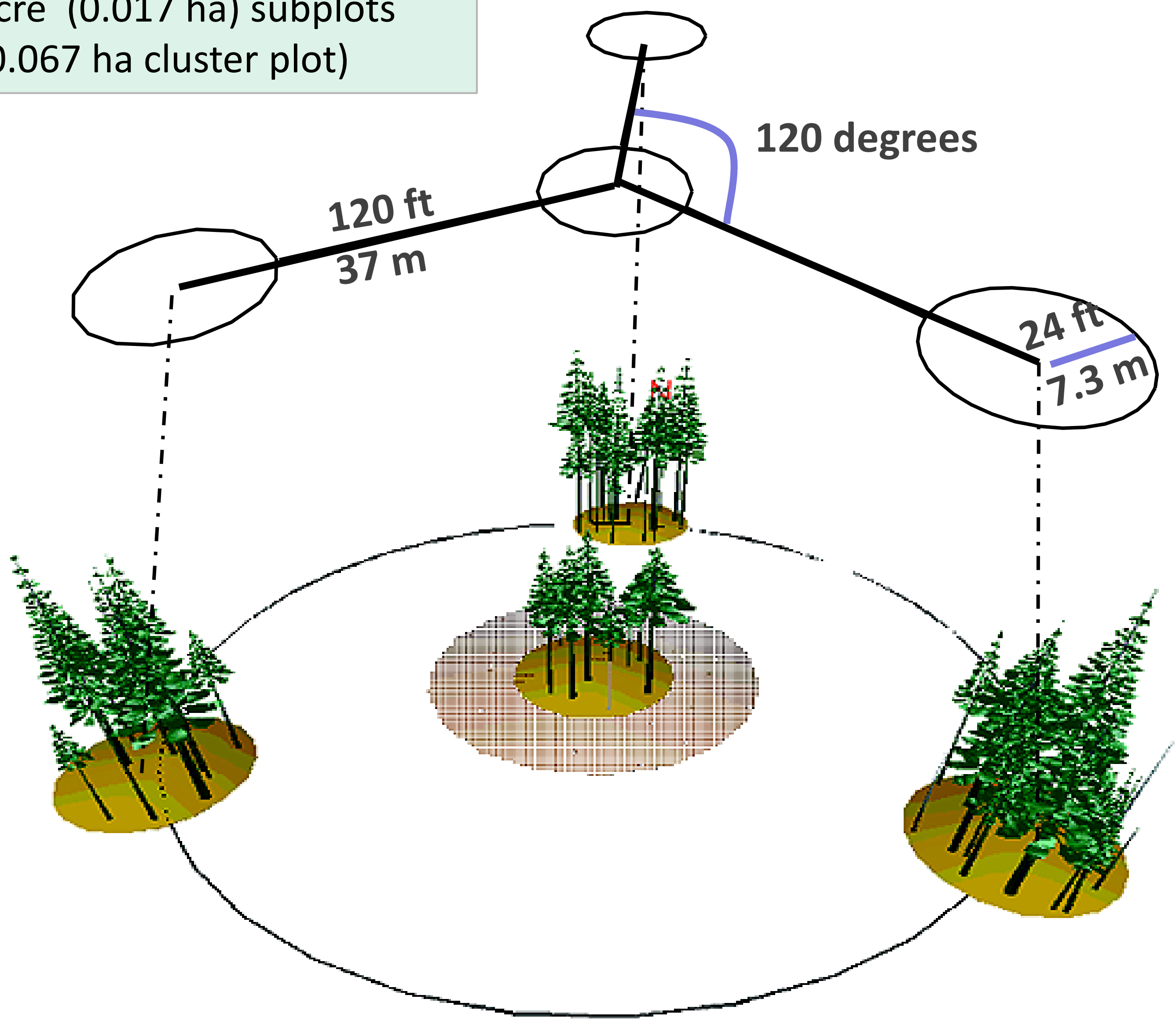
<u>Other</u>
- Invasive plant cover
- Down woody debris
- Soils
- Regeneration

….

<u>Data Processing</u>
- Estimation
- Modelling
- GIS + Remote Sensing
- Reporting

FIA Plot Design

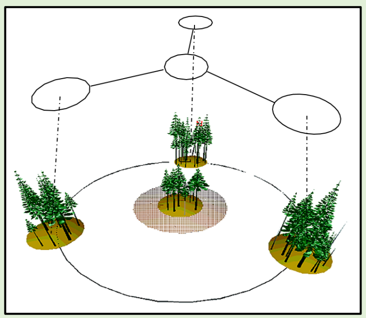
4 X 1/24 acre (0.017 ha) subplots
1/6 acre (0.067 ha cluster plot)



One plot per 6000 acres (2700 ha)

>300,000 plots at full implementation!!

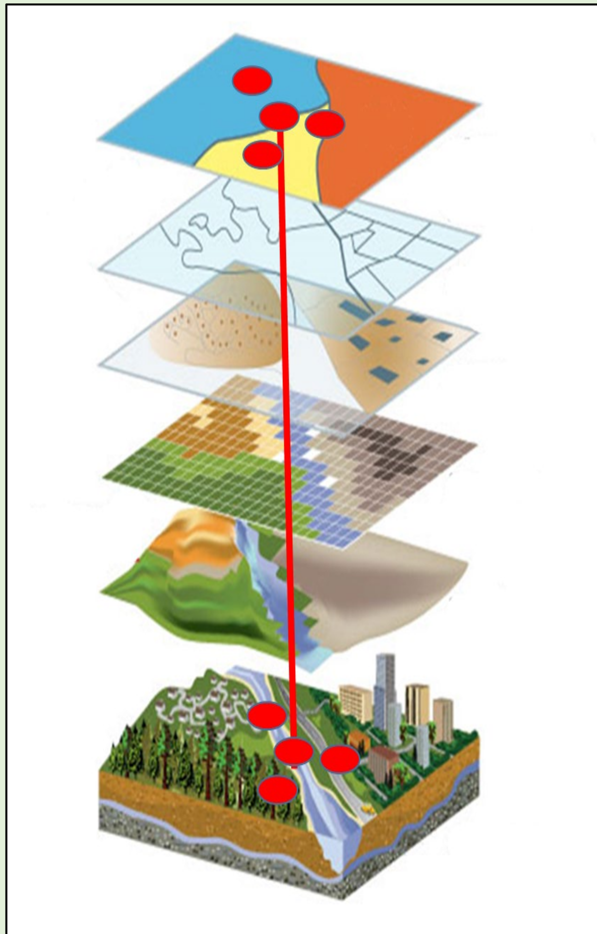
How does FIA use remote sensing and GIS?

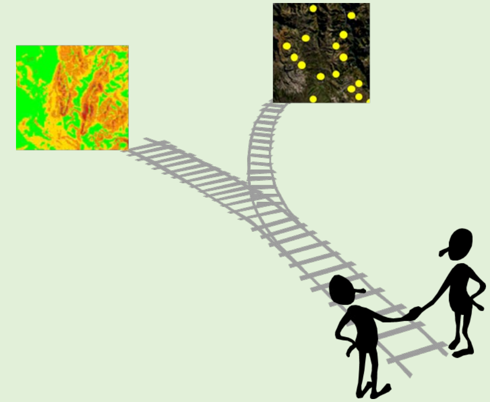


GPS

+

Plots labeled with **GIS and Remote Sensing data**



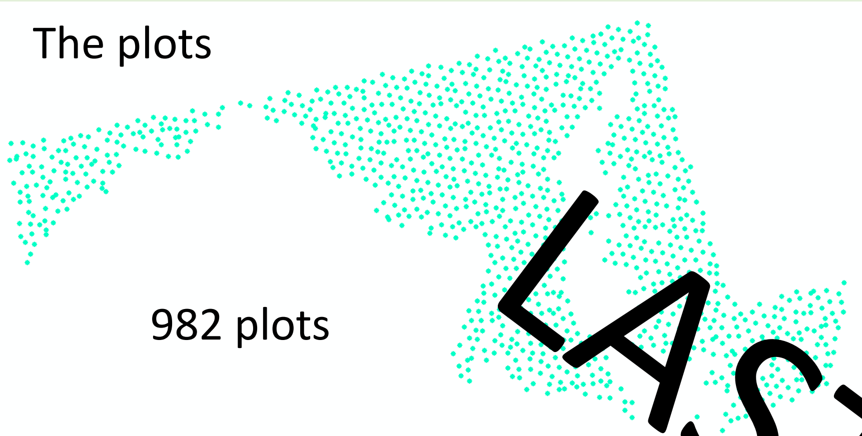


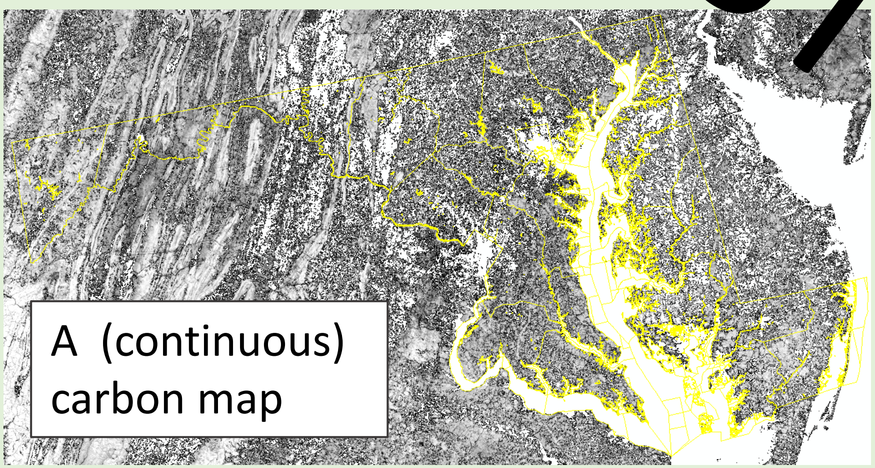
Improved estimation through stratification and other model-assisted estimation methods

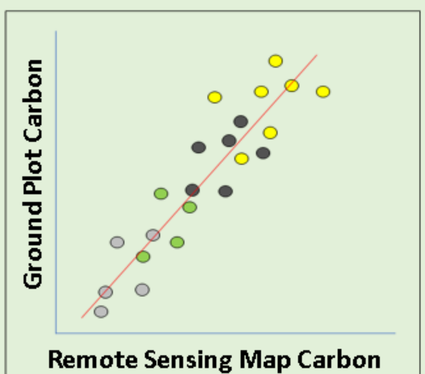
Maps for **model-based estimation**, additional context for tabular estimates

So what did *I* do? *I compared* $v(\bar{Y})$ from MAR with that from PS

LAST YEAR!

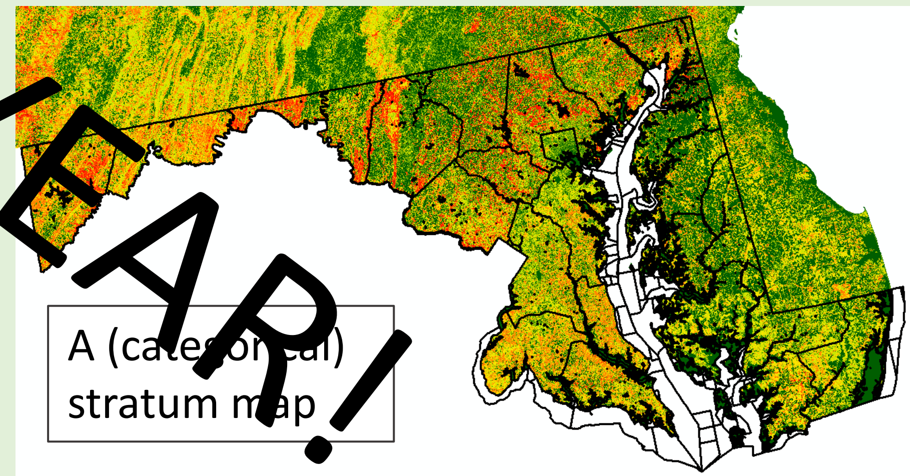






$$\bar{Y}_{MAR} = \bar{y} + b(\bar{X} - \bar{x})$$

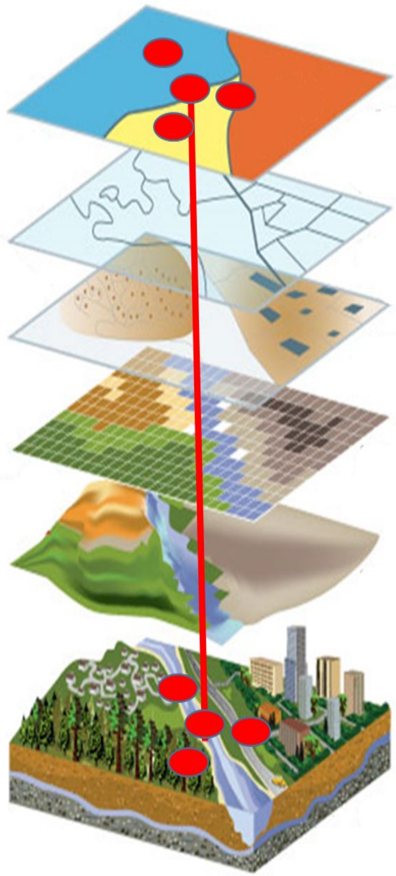
$$v(\bar{Y}_{MAR}) = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n(n-2)}$$



$$\bar{Y}_d = \sum_{h=1}^{H} \frac{N_h}{N} \bar{Y}_{hd}$$

$$v(\hat{Y}_d) = \frac{A_T^2}{n}\left[\sum_{h}^{H} W_h n_h v(\bar{Y}_{hd}) + \sum_{h}^{H}(1 - W_h)\frac{n_h}{n} v(\bar{Y}_{hd})\right]$$

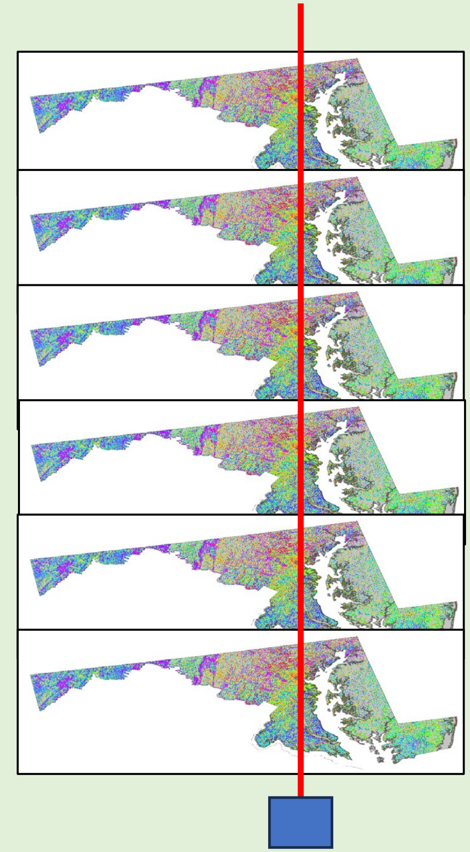
This year: Model-based estimation using remote sensing data.

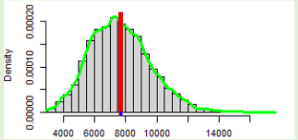


Plots labeled with GIS and Remote Sensing predictors to get a modeling dataset

Forest Carbon = f(predictors)

Modeling dataset subsampled with a Monte Carlo Simulation framework. E.g., subsamples and 1000 MC simulations with Random Forests





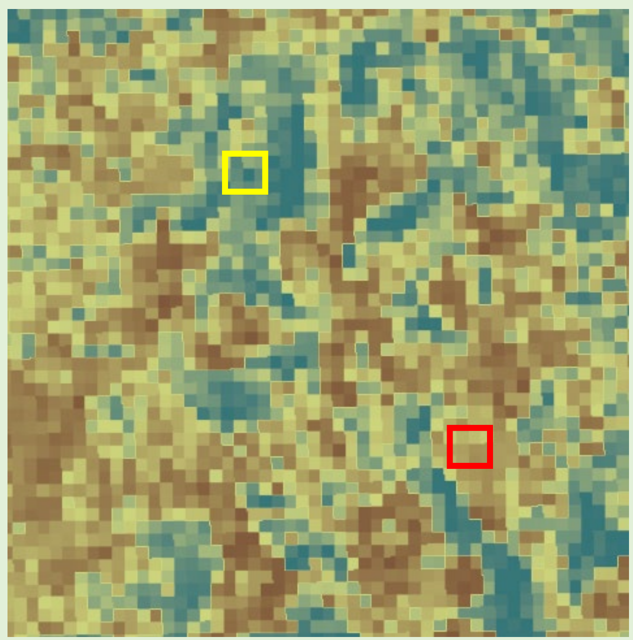
Each pixel in the map has its own mean and variance, i.e. we treat it as N random variables, N= # of pixels in image.

Great! So we can take the mean of the means to estimate the population mean, and average variance to estimate the variance of the mean.

$$\mathrm{Var}\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} \mathrm{Var}(X_i)$$

$\mathrm{Var}(X_i)$ is the variance of the i-th random variable.

BUT! We're forgetting the fact that some of those sets of values used to calculate each pixel's mean are right next to each other, and thus are dependent.



For each MC iteration, the pixels next to each other will likely be more similar to each other than to those far apart.

No problem.. We'll just make a pixel to pixel covariance matrix and fill it in.

10x10 pixel image

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

100x100 pixel covariance matrix!

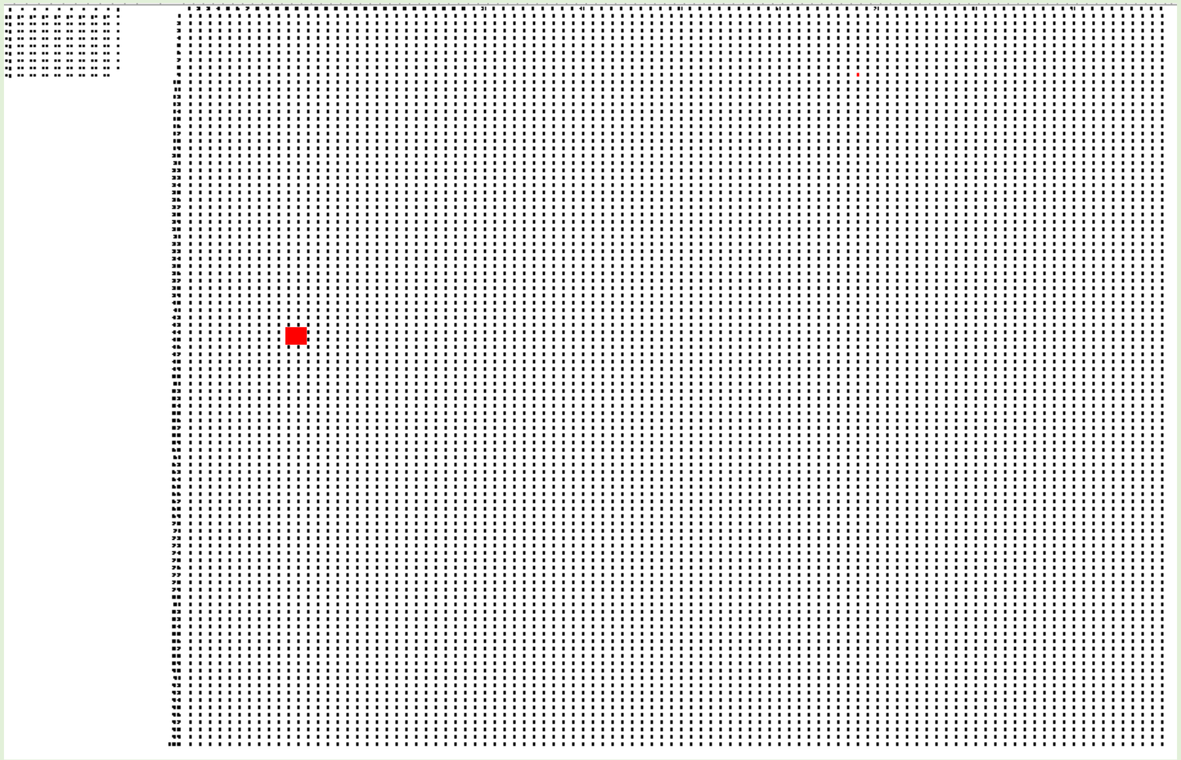
For Maryland, there are 4090 x 2077 100-m pixels or 8,494,9302 = 72,163,835,704,900 covariances to calculate, each with 1000 MC images.

If these were 10m pixels or a larger area, then we have an even bigger problem.

The solution: <u>Sampling</u> the covariance matrix.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Sample pixels 12 and 46 on each of my 1000 MC maps

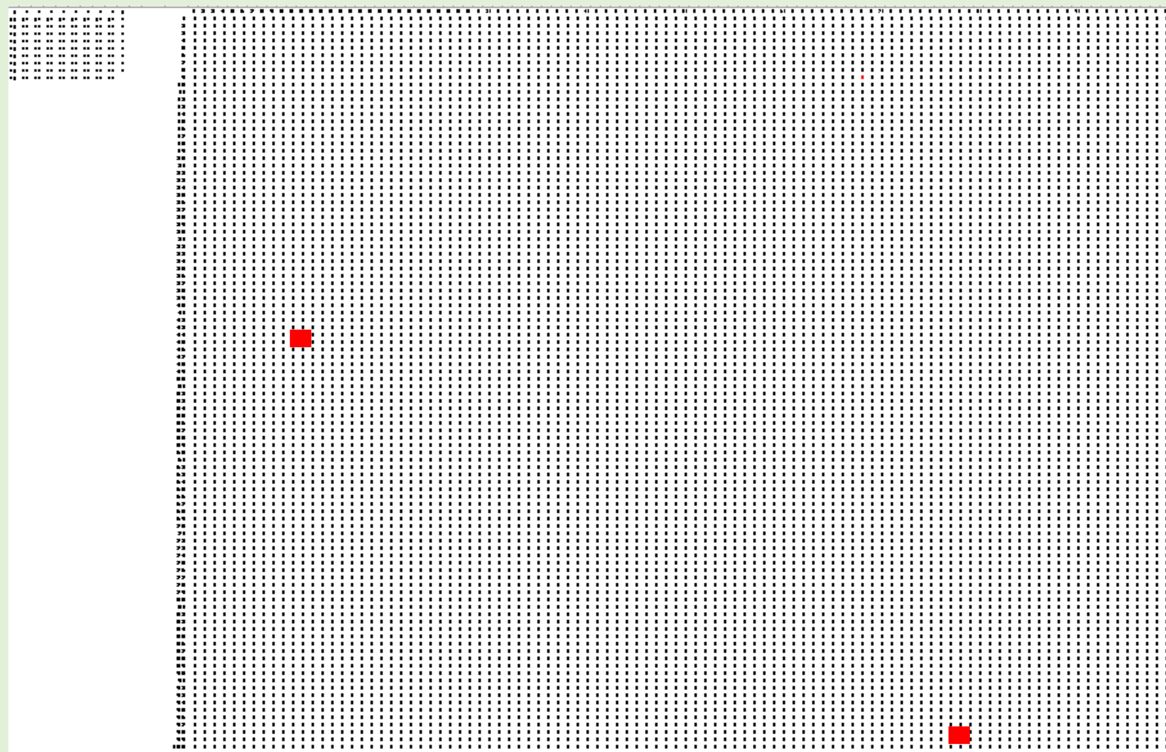


Estimate the covariance of one cell in my covariance matrix…

Take another sample…

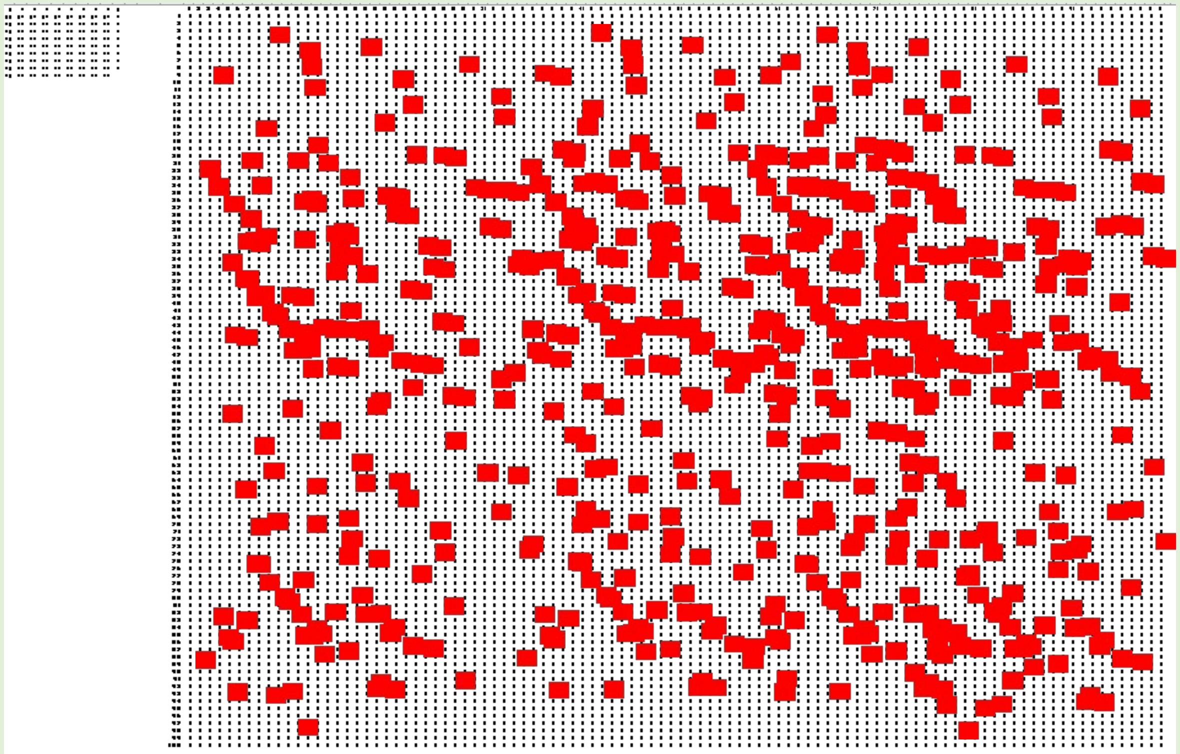
1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Sample pixels 78 and 99 on each of my 1000 MC maps



Estimate the covariance of one cell in my covariance matrix!

Repeat! Take the average of all of the estimated covariances..



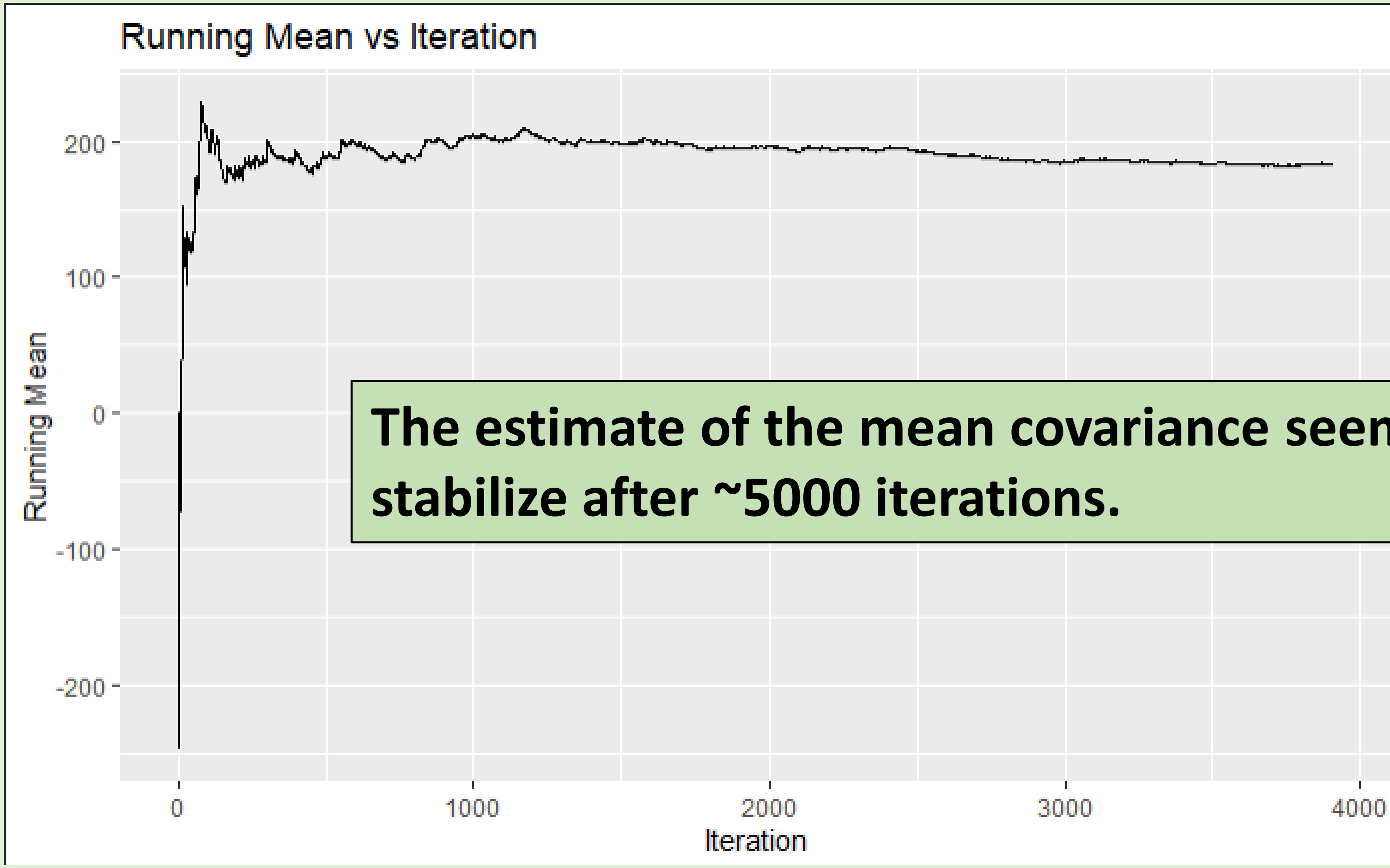
Multiply by the total # of cells in the covariance matrix, and I will have a pretty good estimate of:

$$\mathrm{Var}\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k}\mathrm{Var}(X_i) + \boxed{2\sum_{1\le i<j\le k}\mathrm{Cov}(X_i,X_j)}$$

$\mathrm{Var}(X_i)$ is the variance of the i-th random variable.

$\mathrm{Cov}(X_i,X_j)$ is the covariance between the i-th and j-th random variables

RESULTS:



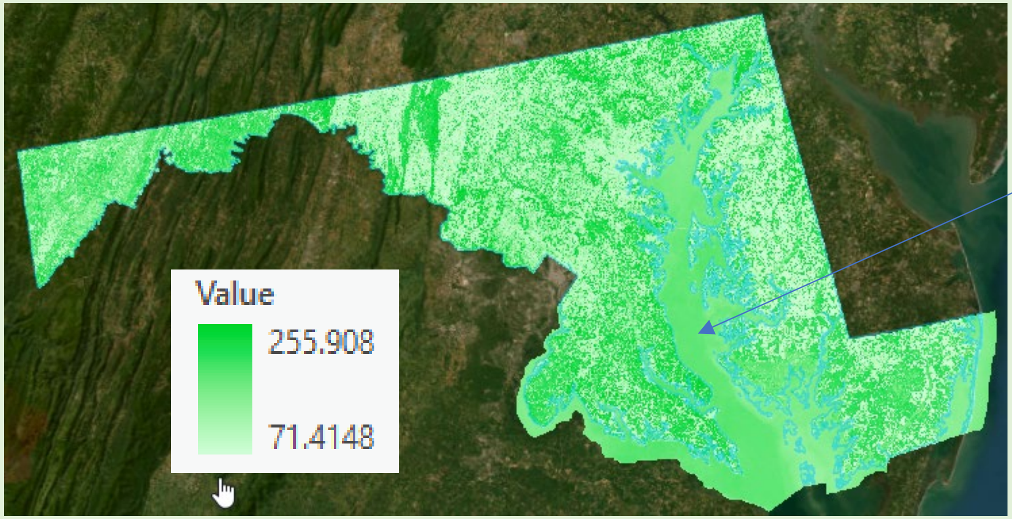
The estimate of the mean covariance seems to stabilize after ~5000 iterations.

$$\text{Std. Error} = \text{SQRT}\left(\sum_{i=1}^{k} \mathrm{Var}(X_i) + 2 \sum_{1 \le i < j \le k} \mathrm{Cov}(X_i, X_j)\right)$$

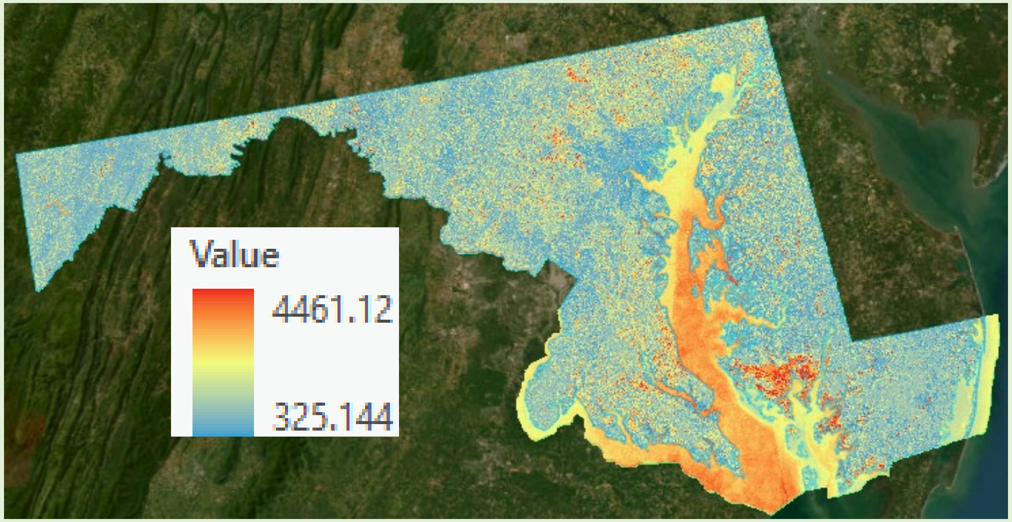
Mean Carbon Mg/ha = 147 +/- **28%** Sampling Error Compare to about **3%** sampling error for the FIA estimates.

Maps:

Pixel mean map



Pixel variance map



Next Steps:

Mask out nonforest areas and compare directly to FIA and UMD carbon maps

Assess assumptions about spatially-dependent covariance of MC simulations

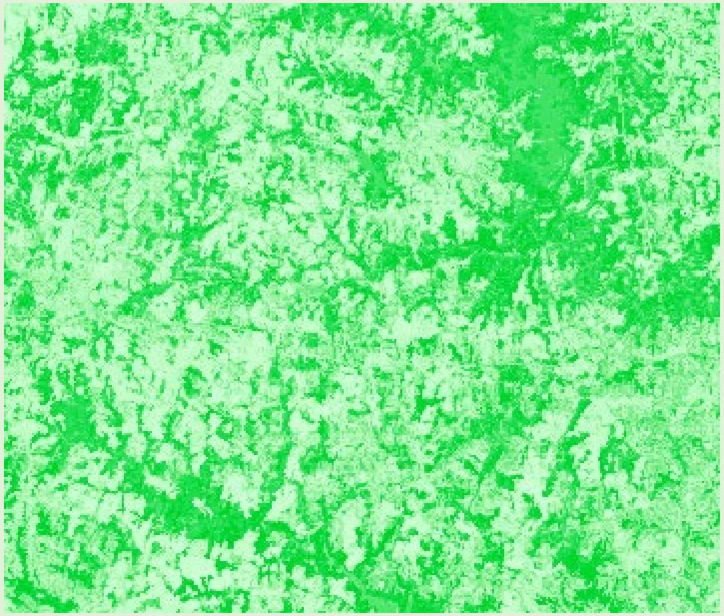
Try to make it faster! In Google Earth Engine, 15 minutes to make 500 maps; 2 hours to get 10,000 random covariance sample points.

Questions and Discussion:

Am I on the right track?

Why is the sampling error so much worse with model-based estimation?

Any other ideas for estimating the sum of the covariance matrix efficiently?



Andrew Lister
andrew.lister@usda.gov