LUTs

Uncertainty and spatial diversity

Javier Pacheco-Labrador

Max Plank Institute for Biogeochemistry

Max Planck Institut
for Biogeochemistry

SENSECO

COST Action CA17134 SENSECO
WG4&1 Training School
November 18 -21 , 2019.
CETAL – INFLPR, Măgurele, România

Layout

1. Look-up tables
   - Concept
   - Methods
2. Uncertainty and spatial variability

1. Introduction to LUTs

- “*array that replaces runtime computation with a simpler array indexing operation*”

X_1	X_2	…	X_n	Y_1	Y_2	…	Y_m
1	2	8	32	5	3	74	1
74	3	28	3	9	9	4	4
5	36	38	69	3	9	3	4
6	4	13	9	[illegible]	9	14	36
74	66	96	16	47	4	5	51
6	8	6	198	6	1	75	1
6	4	6	9	1	97	6	7

1. Introduction to LUTs

- RTM $\{X_1, X_2, \ldots X_n\} \xrightarrow{f_{\text{slow}}\,(X_1,X_2,\ldots X_n)} \{Y_1, Y_2, \ldots Y_n\}$

- In Remote Sensing
  - Traditionally used to invert RTM

$$\{X_1, X_2, \ldots X_n\} \xrightarrow{find\left(X_{\mathrm{p}} \middle| Y_{\mathrm{p}} \sim Y_{\mathrm{obs}}\right)} \{Y_1, Y_2, \ldots Y_n\}$$

  - Train statistical predictive models (Hybrid models)

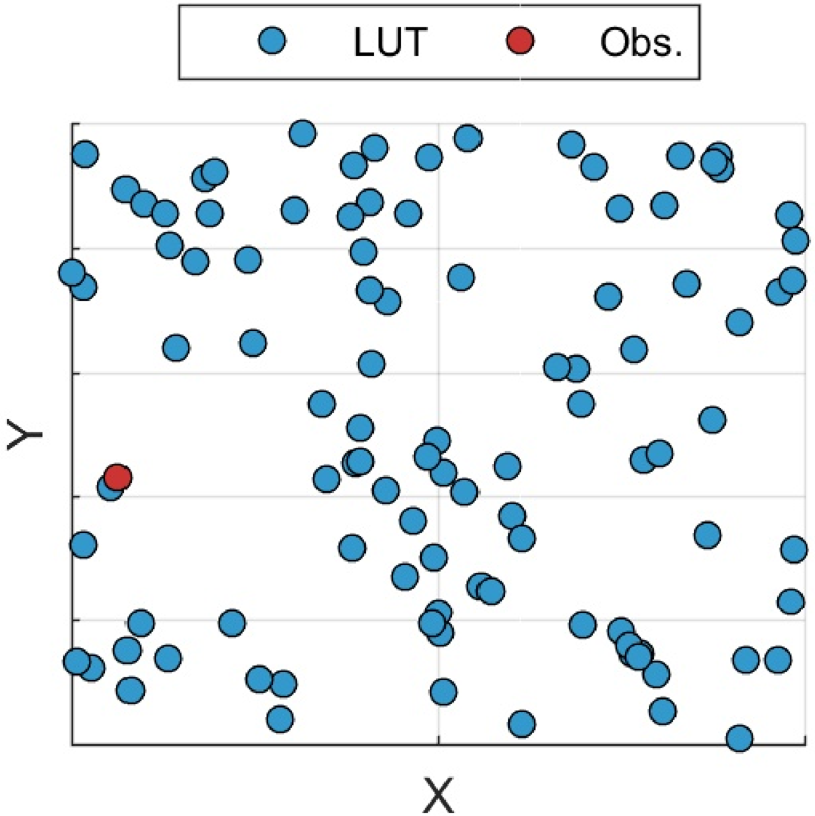
$$\{X_1, X_2, \ldots X_n\} \xrightarrow{f_{\text{fast}}\,(X_1,X_2,\ldots X_n)} \{Y_1, Y_2, \ldots Y_n\} \quad \text{Forward}$$

$$\{Y_1, Y_2, \ldots Y_n\} \xrightarrow{f_{\text{fast}}\,(Y_1,Y_2,\ldots Y_n)} \{X_1, X_2, \ldots X_n\} \quad \text{Inverse}$$

1. Introduction to LUTs

- Coverage of variable space

Uniform random distribution

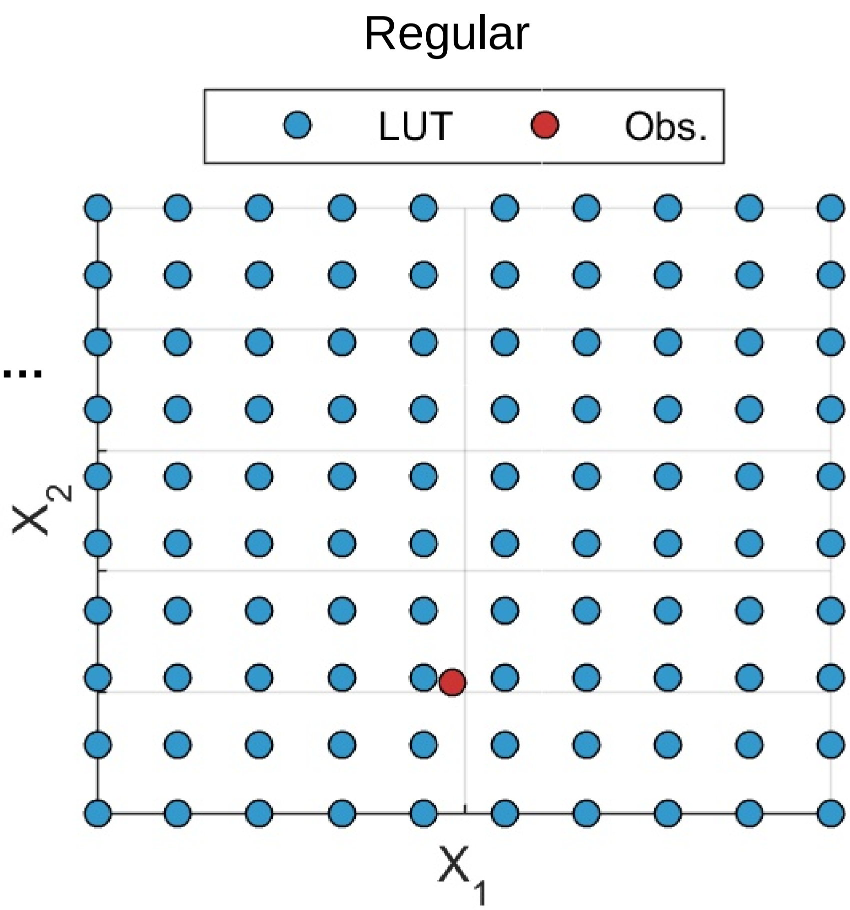



```
n = 100;

X1 = rand([n, 1]);
X2 = rand([n, 1]);
```

1. Introduction to LUTs

- Coverage of variable space

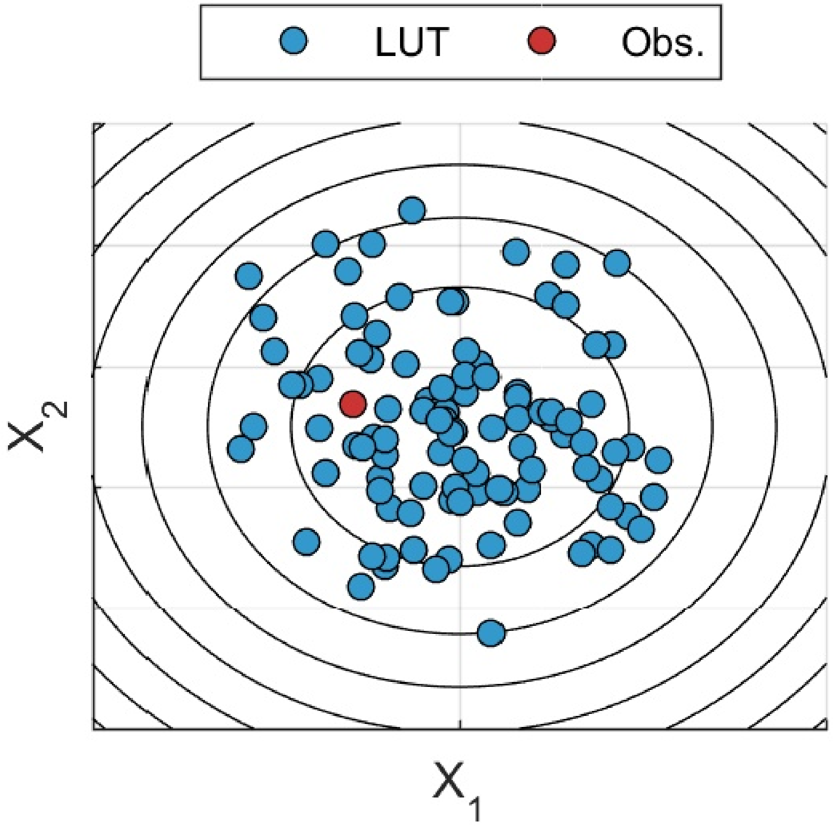
```
n = 100;
nr = round(sqrt(n));

[x1, x2] = ...
    meshgrid(linspace(0, 1, nr), ...
    linspace(0, 1, nr));
```



1. Introduction to LUTs

- Coverage of variable space

Uncorrelated multivariate




```
n = 100;
mu = .5
sigma = .15

X1 = normrnd(mu, sigma, [n,1]);
X2 = normrnd(mu, sigma, [n,1]);
```

1. Introduction to LUTs

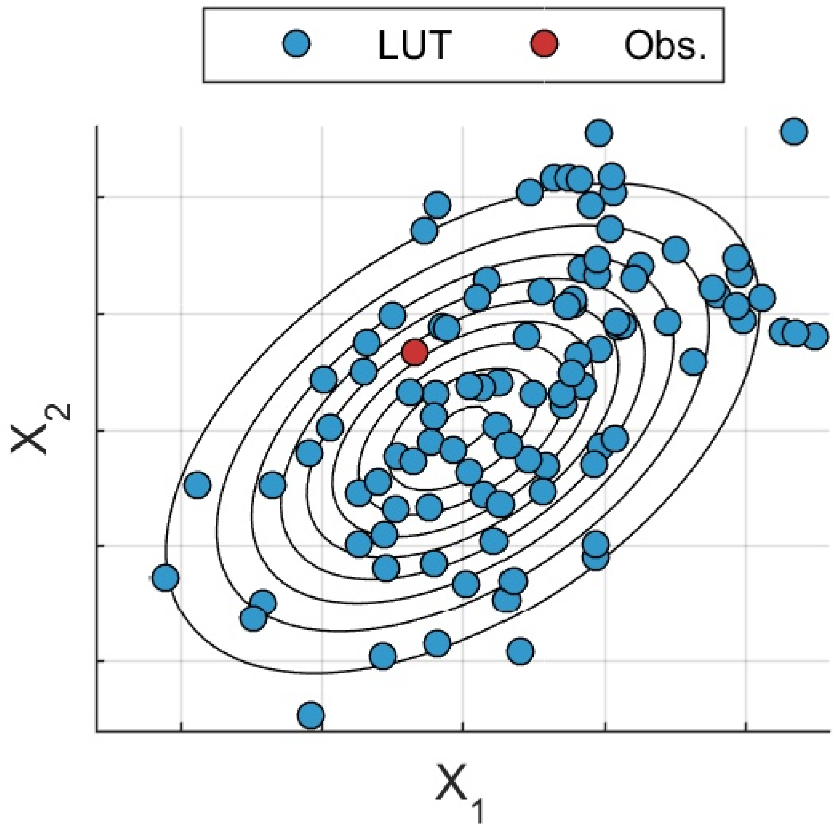
- Coverage of variable space

```
n = 100;
mu = [.5 .5];
Sigma = [1 .5;.5 1];

X = mvnrnd(mu, Sigma, n);
X1 = X(:,1); X2 = X(:,2);
```

Uncorrelated case…

```
n = 100;
mu = [.5 .5];
Sigma = [1 0;0 1];

X = mvnrnd(mu, Sigma, n);
```

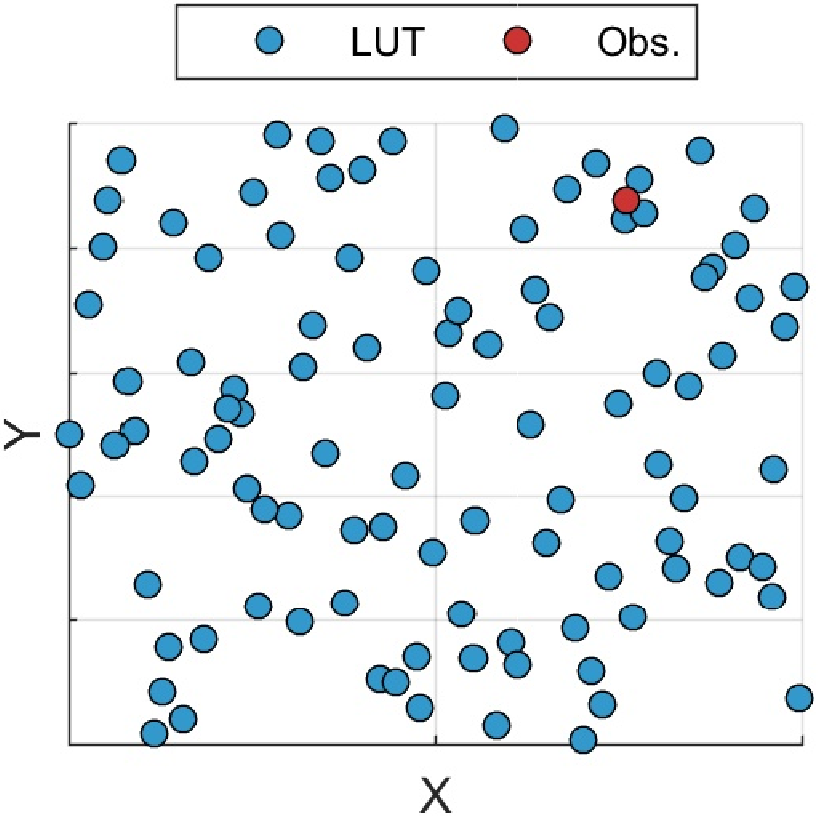
Correlated multivariate



1. Introduction to LUTs

- Coverage of variable space

Latin hypercube sampling



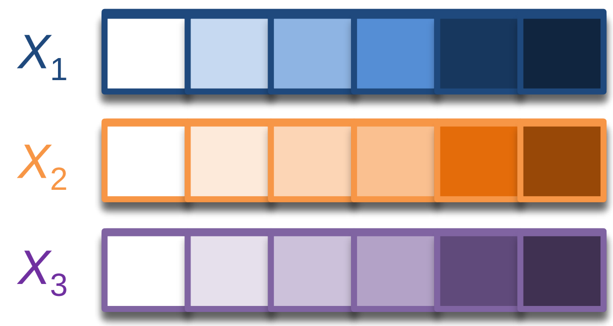

```
n = 100;

X = lhsdesign([n,2]);
X1 = X(:,1);  X2 = X(:,2);
```

1. Introduction to LUTs

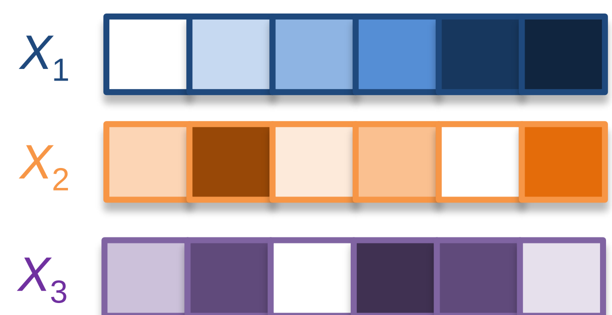
▪ Coverage of variable space

LHS optimizes coverage for a given *n* size:

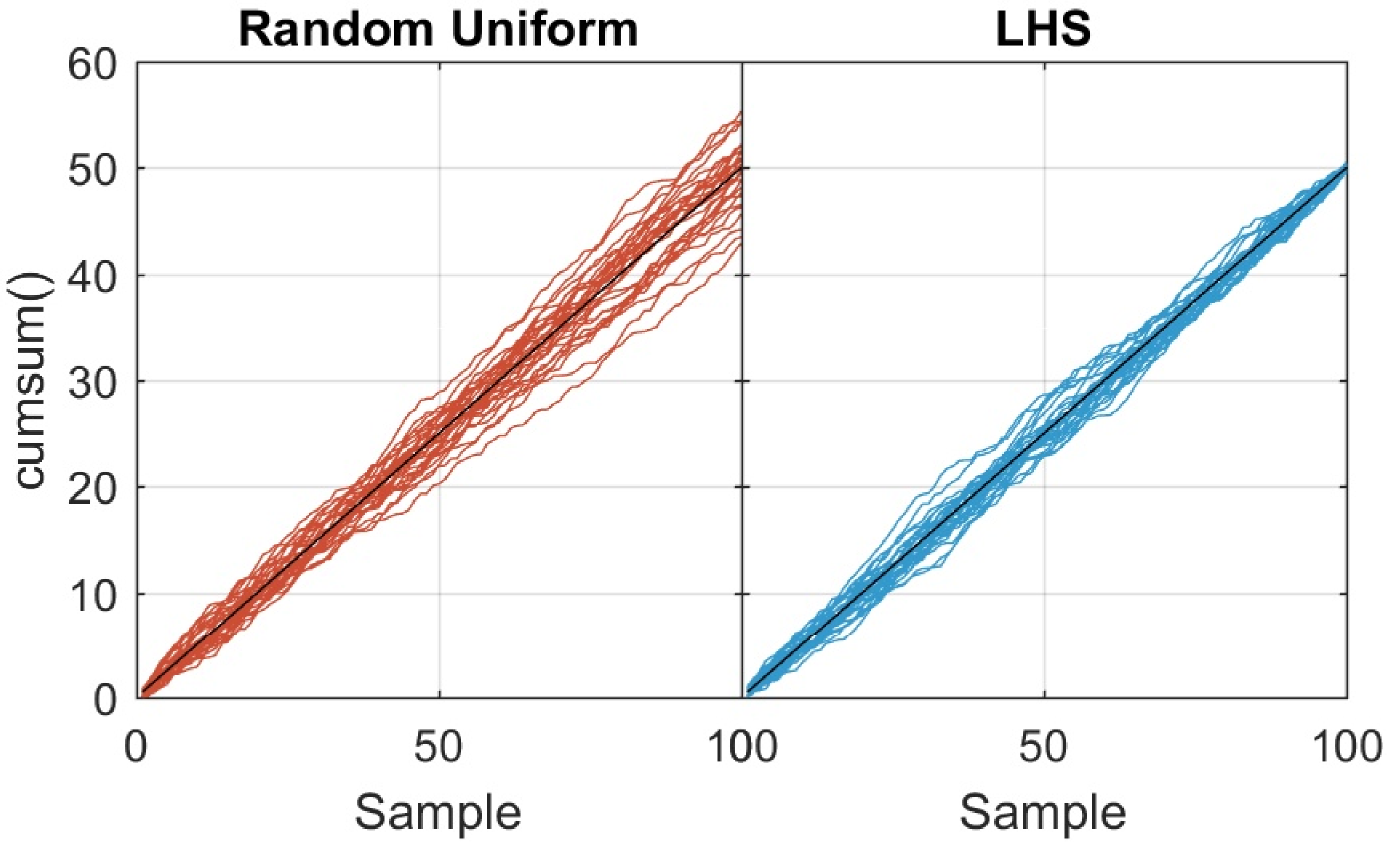
- Split cumulative density function into *n* disjoint intervals of equal probability
- One value is selected at random from each interval
- X1 paired at random with X2
- [X1,X2] paired at random with X3
- …
- Additional criteria can be iteratively optimized (e.g. minimize correlation, maximize distances…)



`While ~criteria OR iter<itmax`



1. Introduction to LUTs



1. Introduction to LUTs

- Combining LHS & Conditional Probability
  - The joint probability / dependence of some of the parameters (X) is known for some of the parameters.

    - LHS to simulate parameters with no prior information about their dependence on other parameters (X_{ind})

    - Use prior knowledge to define values of parameters that depend on other parameters ($X_{\mathrm{dep}} = f(X_{\mathrm{ind}}) + \varepsilon$)

Reduce / Avoid unrealistic combinations

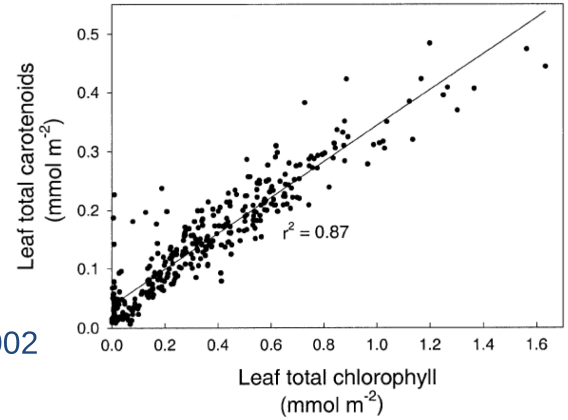
1. Introduction to LUTs

- Combining LHS & Conditional Probability
  - Example: PROSPECT model

LHS

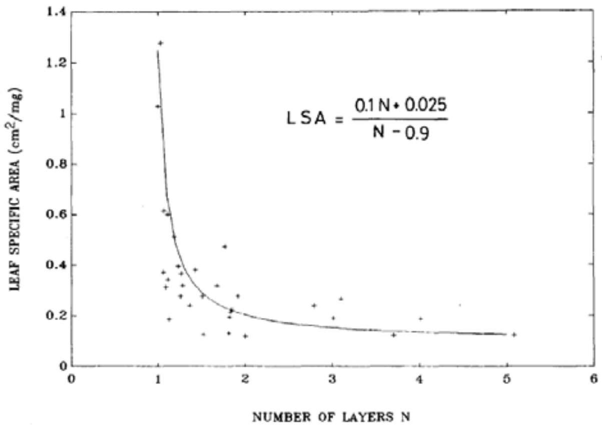
C_{ab}	C_{ar}	C_m	C_w	N
20.3	6.8	0.0019	0.018	1.65
40.5	14.5	0.0025	0.0090	1.30
...	...	...	...	...

$C_{ar} = f(C_{ab}) + \mathcal{N}(0,\sigma)$



Sims and Gamon, 2002

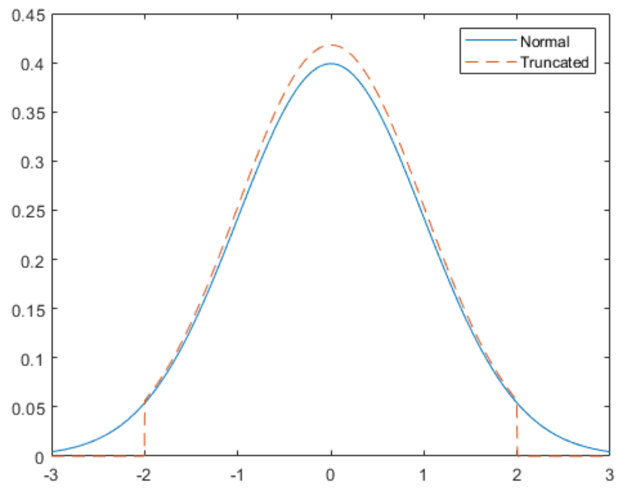
$N = f(C_m) + \mathcal{N}(0,\sigma(f(C_m)))$

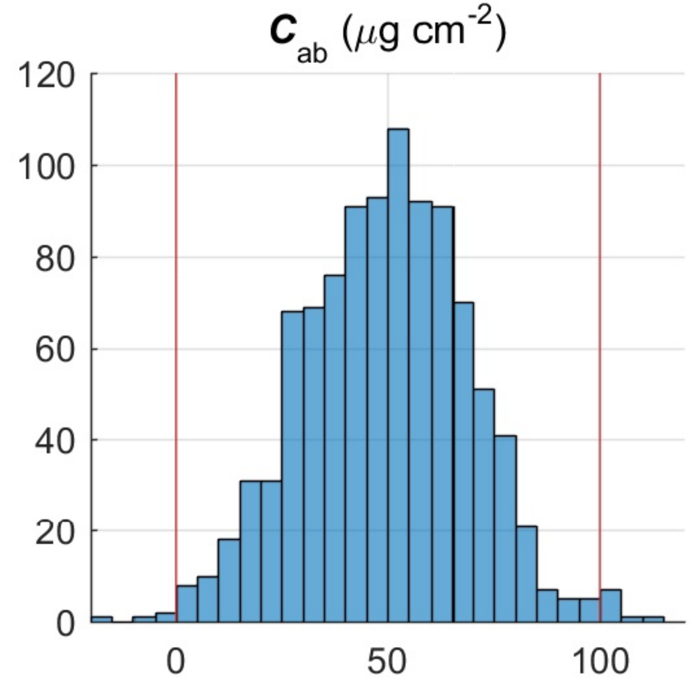


Jacquemoud and Baret, 1990

1. Introduction to LUTs

- Truncated distributions
  - Prevent unrealistic / physically impossible values
  - Truncation modifies PDF





1. Introduction to LUTs

- Look for methods already developped
  - Botev, Z.I., & Ecuyer, P.L. (2015). Efficient probability estimation and simulation of the truncated multivariate student-t distribution. In, 2015 Winter Simulation Conference (WSC) (pp. 380-391)
  - Botev, Z.I. (2017). The normal law under linear restrictions: simulation and estimation via minimax tilting. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79, 125-148

1. Introduction to LUTs

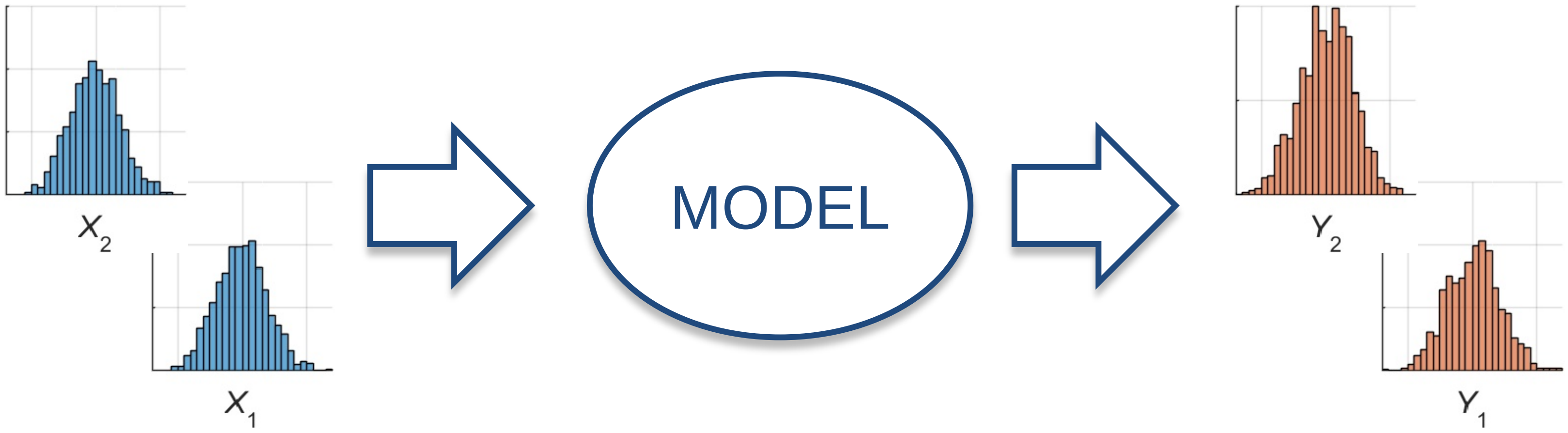
- Alternatively, iteratively sample values within ranges

  - While any(I)
    - Sample from the distributions $x = \mathcal{N}(\text{params})$
    - $I = x < \text{LB} \mid x > \text{UB}$
    - Keep results within range (isfalse(I))

  - Can be inefficient!!

2. Uncertainty and spatial variability

- The approaches seen for LUT generation are not but ways to design a Monte Carlo scheme for run a model over a given distribution of parameters
  - This can be used to propagate uncertainties
  - Can be also used to estimate the effect of parameters' variability in a integrated signal (e.g., remote sensing pixel)



2. Uncertainty and spatial variability

- Considerations
  - Covariance of uncertainties
    - e.g., dark current and instrument sensitivity are function of temperature
    - e.g., chlorophyll and carotenoids are correlated; and can correlate with leaf area or water content in some ecosystems
  - Realistic ranges
    - Physically plausible
    - Instrument / Ecosystem / time specific
  - If the LUT will be / train an estimator, is better envelope the expected ranges. Avoid extrapolation

THANKS!