Optimal Sampling Strategies

An Application of Graph Theory

Joel A. Gongora
Dr. H.P. Marshall
Dr. Nancy F. Glenn
Dr. Inaac Senocak

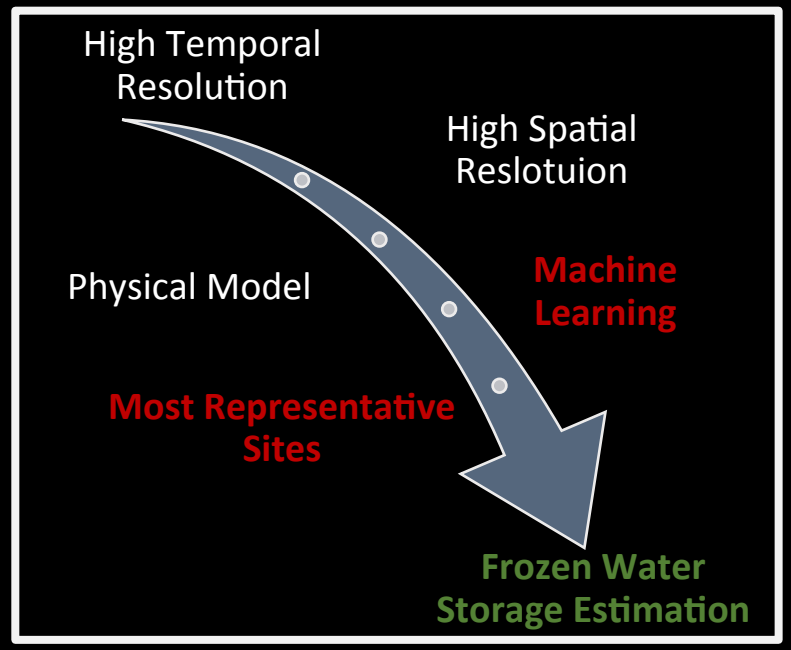
Optimal Sampling Strategies

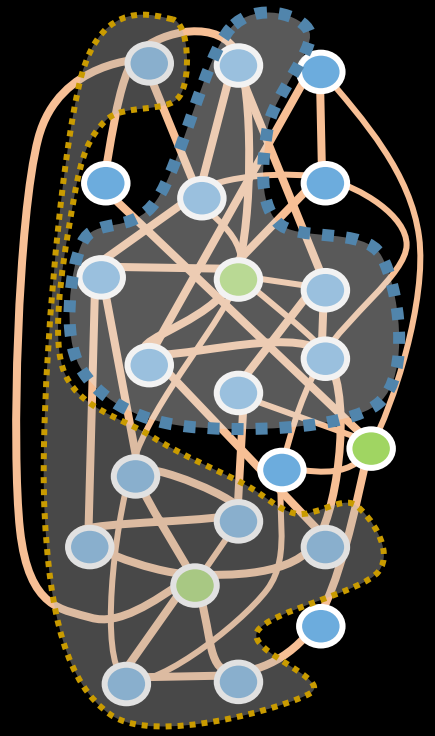
An Application of Graph Theory

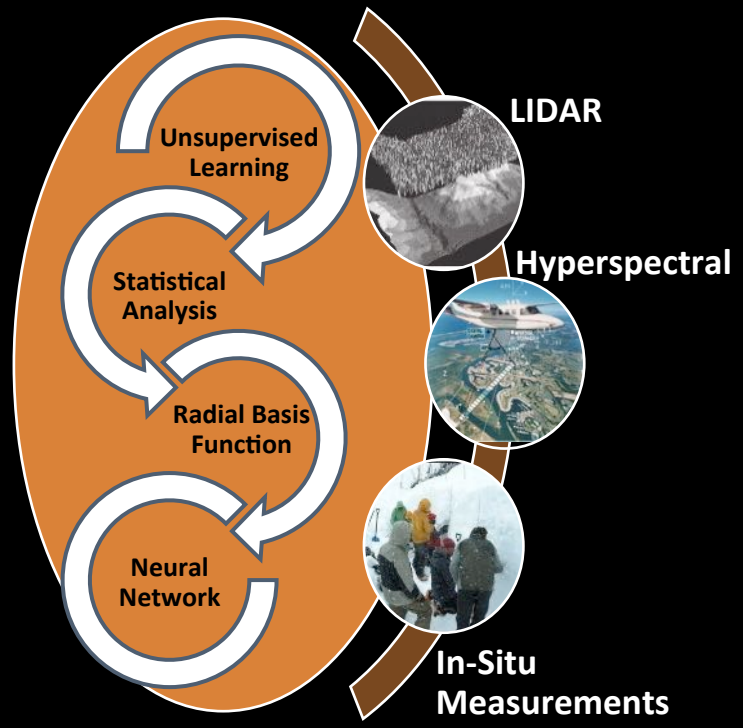
We thank the NASA CLPX and SnowEx project. This work was funded by a NASA EPSCor grant, Monitoring Earth's Hydrosphere: Integrating Remote Sensing, Modeling, and Verification.

Over Arching Goal

Develop a machine learning approach capable of merging a variety of datasets to identify locations ideal for in-situ measurements of snow. The results of this work will be an improvement in water supply estimates and will allow for detection of key validation points for current and future run-off models.







Overview of Neural Network Training

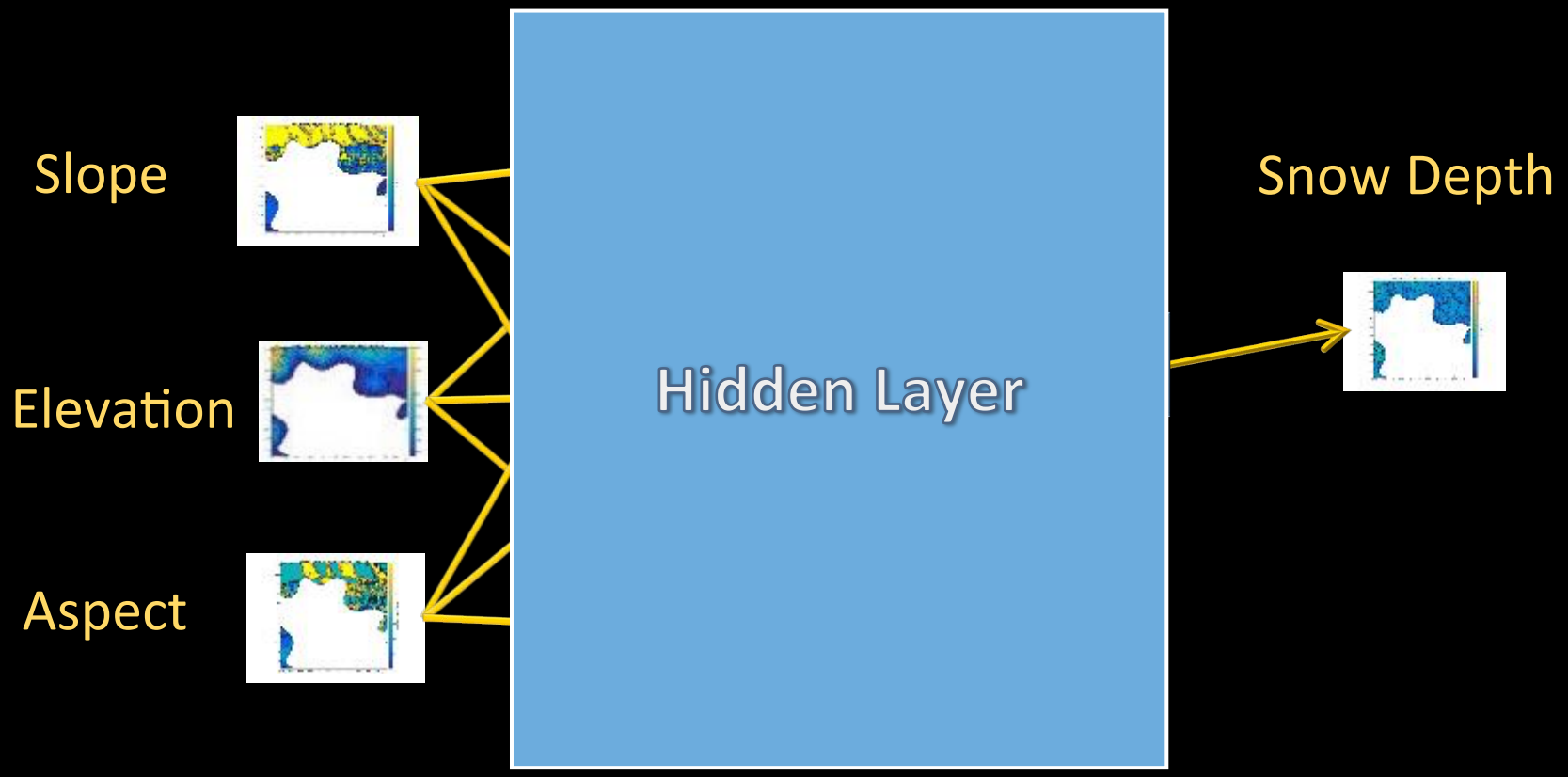
Outline

- Radial Basis Function Neural Network
- Partitioning
- What is Graph Theory?
- Defining Graphs.
- Hubs/Modularity
- Optimization
- Some Results

BOISE STATE UNIVERSITY

cryo GARS CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

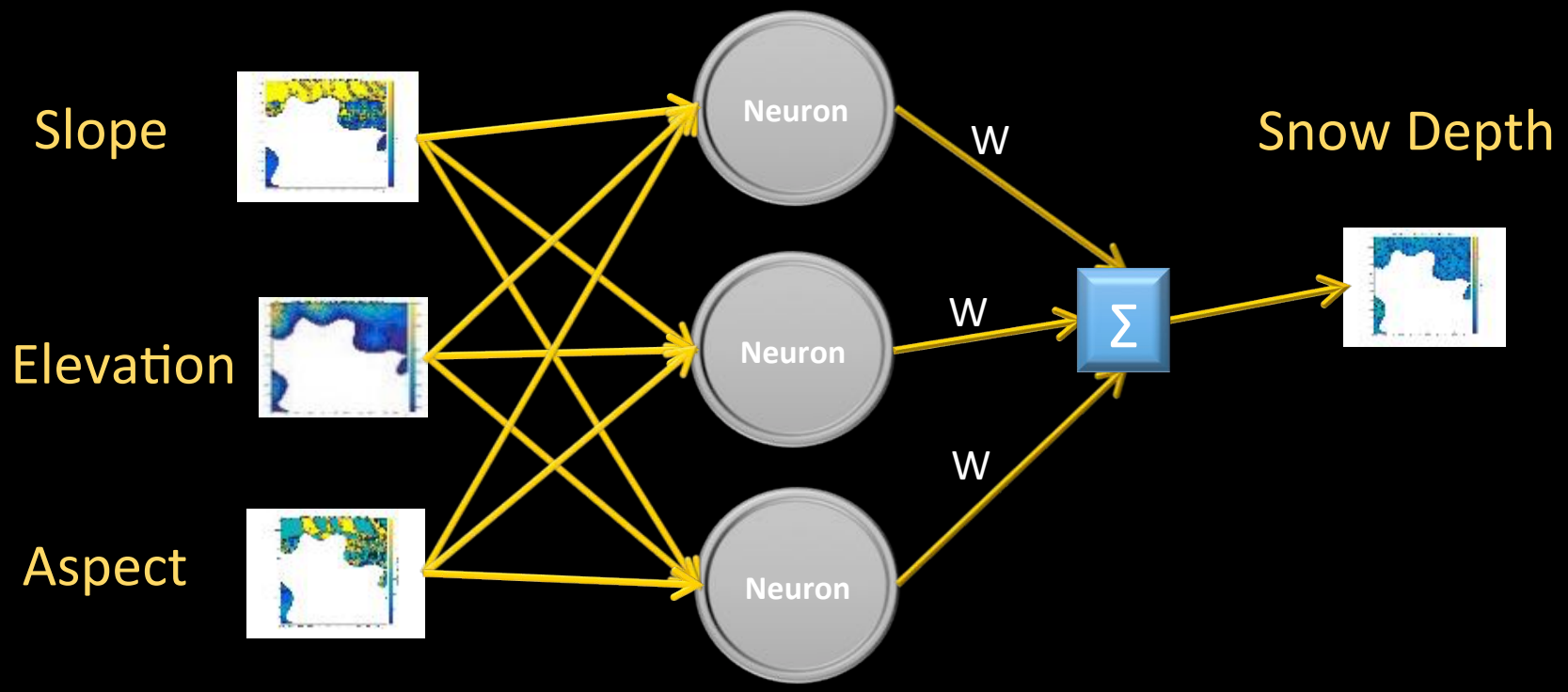
Single Layer Neural Network



BOISE STATE UNIVERSITY

cryo GARS CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

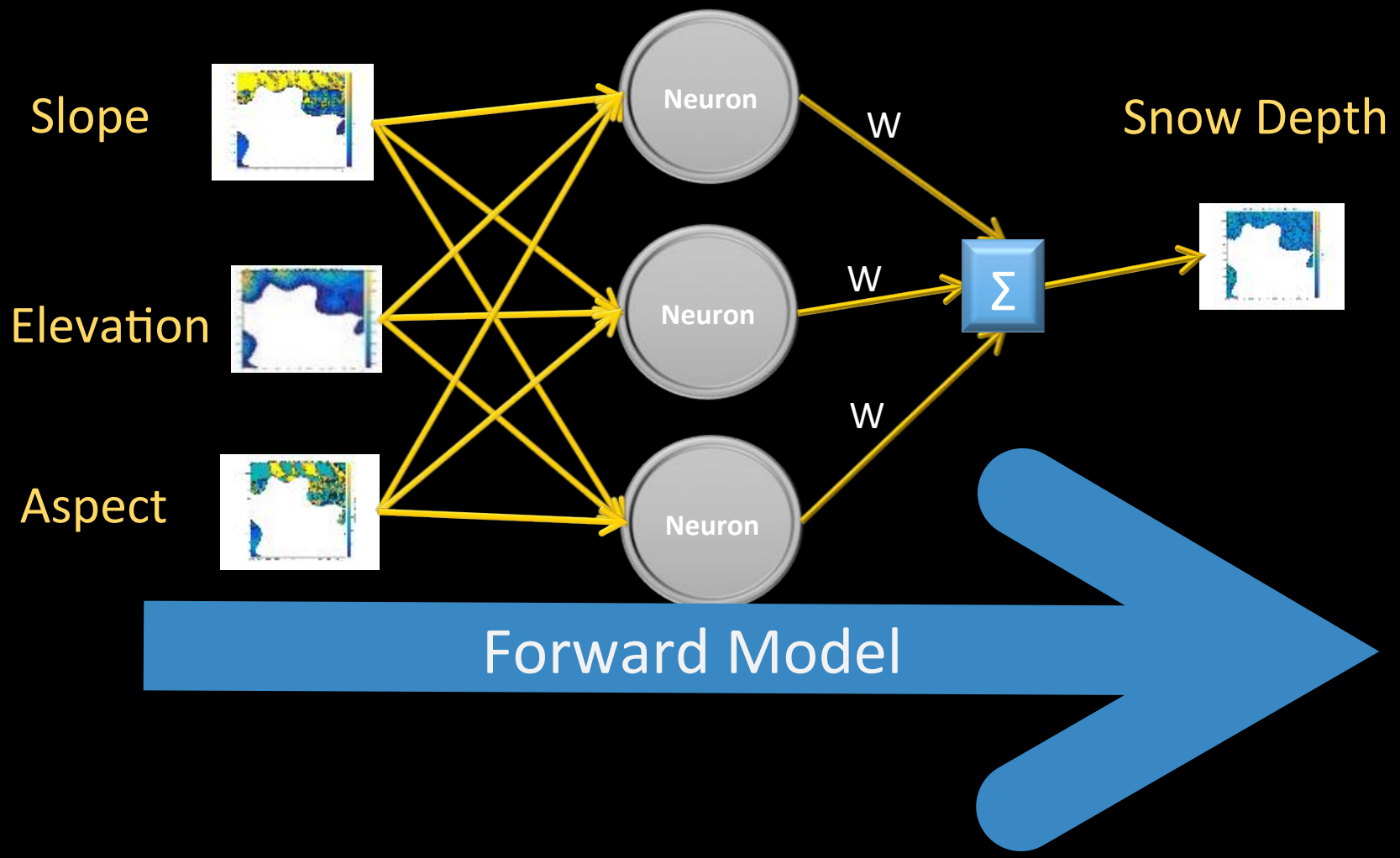
Single Layer Neural Network



BOISE STATE UNIVERSITY

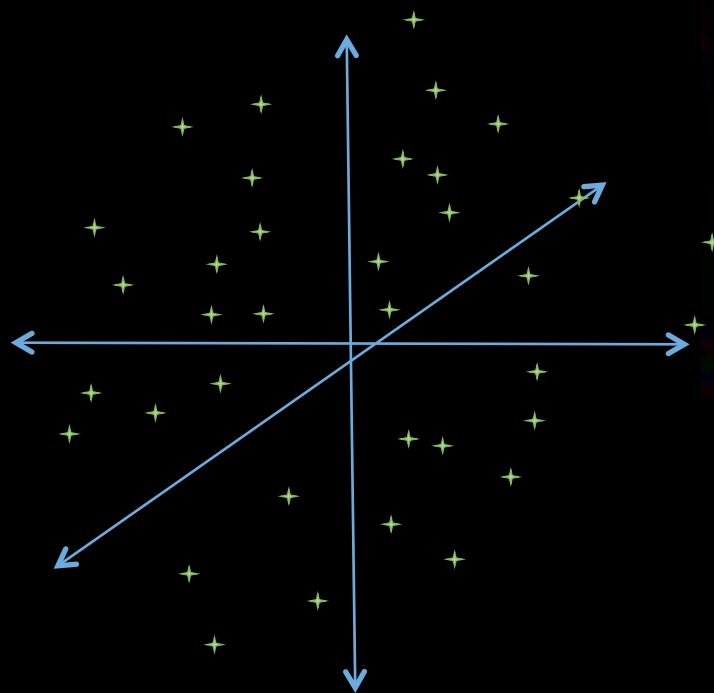
cryo GARS CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

Single Layer Neural Network



Partitioning Data

Feature Matrix



Entire Data Set

Slope 1	Elev 1	Aspect 1
Slope 2	Elev 2	Aspect 2
Slope 3	Elev 3	Aspect 3
Slope 4	Elev 4	Aspect 4
Slope 5	Elev 5	Aspect 5
.	.	.
.	.	.
.	.	.
.	.	.

Training Set

Slope 1	Elev 1	Aspect 1
Slope 2	Elev 2	Aspect 2
Slope 3	Elev 3	Aspect 3
Slope 4	Elev 4	Aspect 4
Slope 5	Elev 5	Aspect 5
.	.	.
.	.	.
.	.	.
.	.	.

Cross-Validation Set

Slope 1	Elev 1	Aspect 1
Slope 2	Elev 2	Aspect 2
Slope 3	Elev 3	Aspect 3
Slope 4	Elev 4	Aspect 4
Slope 5	Elev 5	Aspect 5
.	.	.
.	.	.
.	.	.
.	.	.

Testing Set

Slope 1	Elev 1	Aspect 1
Slope 2	Elev 2	Aspect 2
Slope 3	Elev 3	Aspect 3
Slope 4	Elev 4	Aspect 4
Slope 5	Elev 5	Aspect 5
.	.	.
.	.	.
.	.	.
.	.	.

The Approach

Graph Theory + Neural Networks

What is Graph Theory?

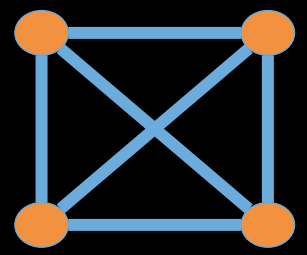
- Relationships defined by connections known as edges.

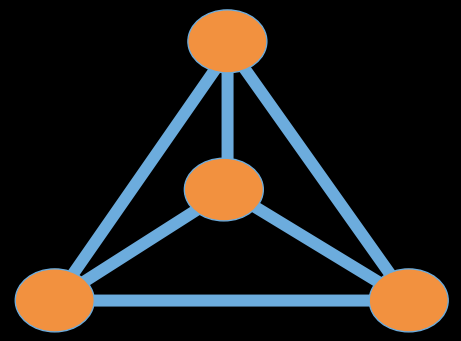
What is Graph Theory?

- Relationships defined by connections known as edges.
- Relationships between individual objects or measures called nodes can then be formed.

What is Graph Theory?

- Relationships defined by connections known as edges.
- Relationships between individual objects or measures called nodes can then be formed.
- A collection of edges and nodes is called a Graph.



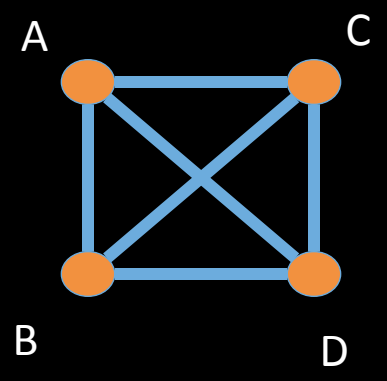


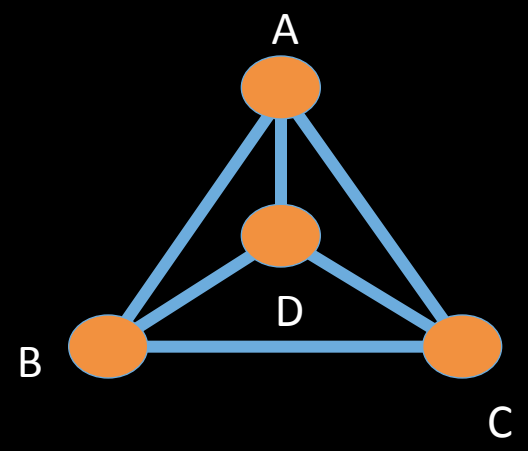
What is Graph Theory?

- Adjacency Matrix

$$A_{ij} =$$

	A	B	C	D
A	0	1	1	1
B	1	0	1	1
C	1	1	0	1
D	1	1	1	0



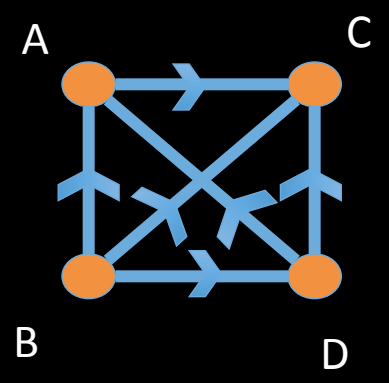


What is Graph Theory?

- Adjacency Matrix

$$A_{ij} =$$

	A	B	C	D
A	0	-1	1	-1
B	1	0	1	1
C	-1	-1	0	-1
D	1	-1	1	0

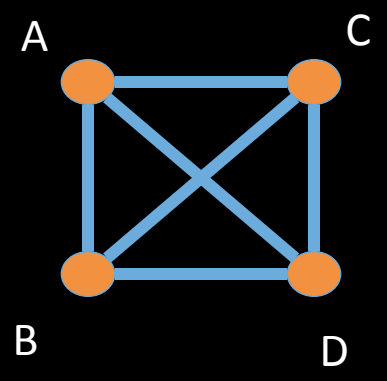


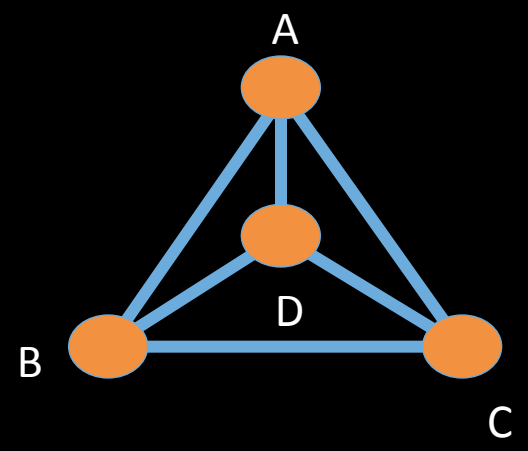
What is Graph Theory?

- Degree Matrix

$$D_{ii} =$$

	A	B	C	D
A	3	0	0	0
B	0	3	0	0
C	0	0	3	0
D	0	0	0	3



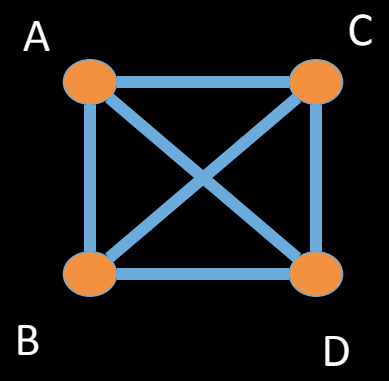


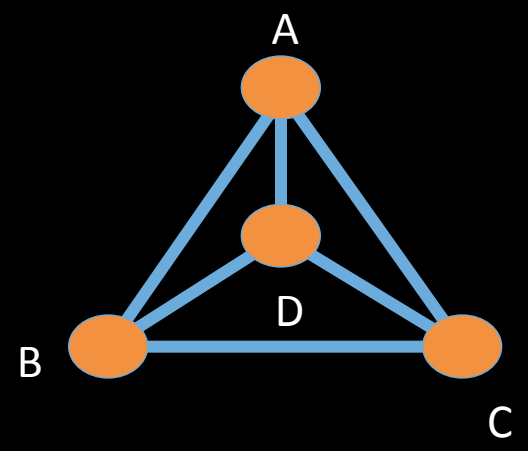
What is Graph Theory?

- Laplacian Matrix

$$D_{ii} - A_{ij} = L_{ij} =$$

	A	B	C	D
A	3	-1	-1	-1
B	-1	3	-1	-1
C	-1	-1	3	-1
D	-1	-1	-1	3





Graph Properties

Number of Nodes/Links

Directed or Simple

Degree Distribution

Diameter

Cluster Analysis

Spanning Trees

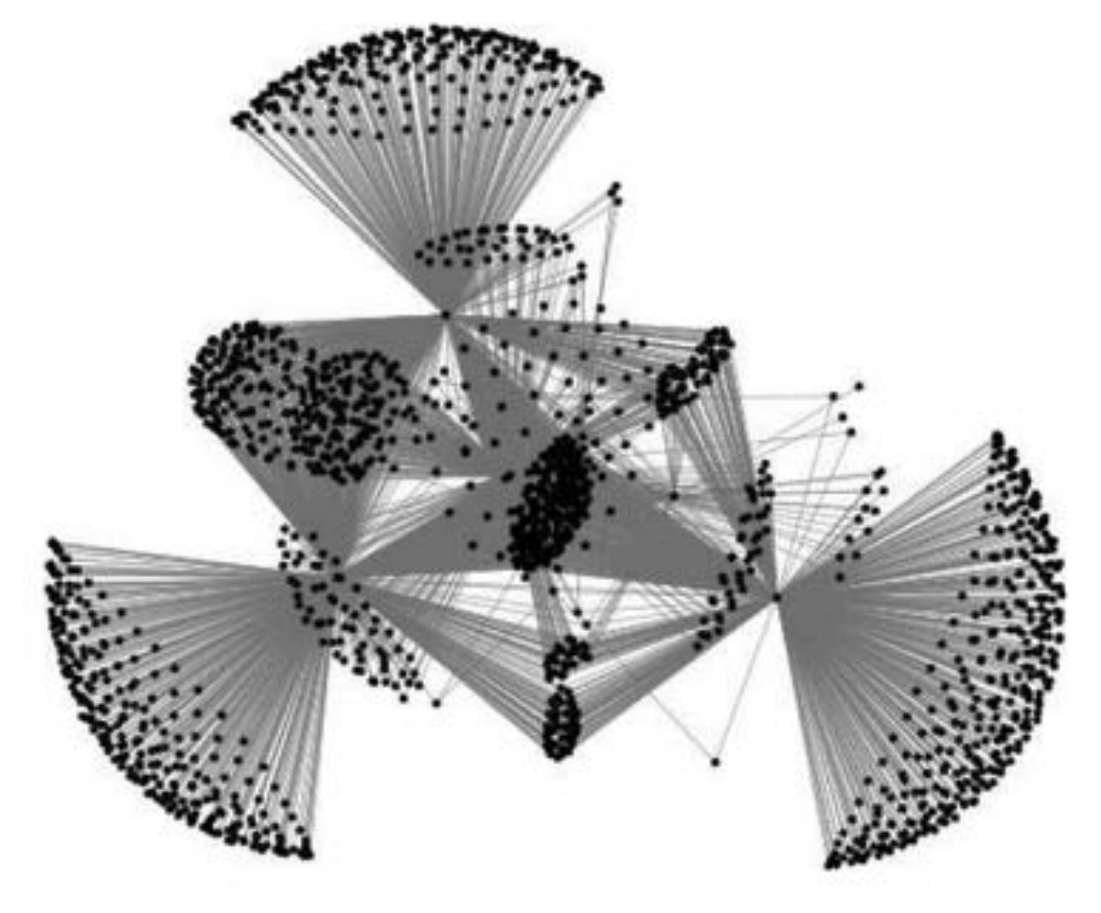
Hamiltonian Cycles

Shortest Path

Hubs/Centrality

BOISE STATE UNIVERSITY

cryo GARS
CRYOSPHERE GEOPHYSICS AND REMOTE SENSING
earth.boisestate.edu/cryogars

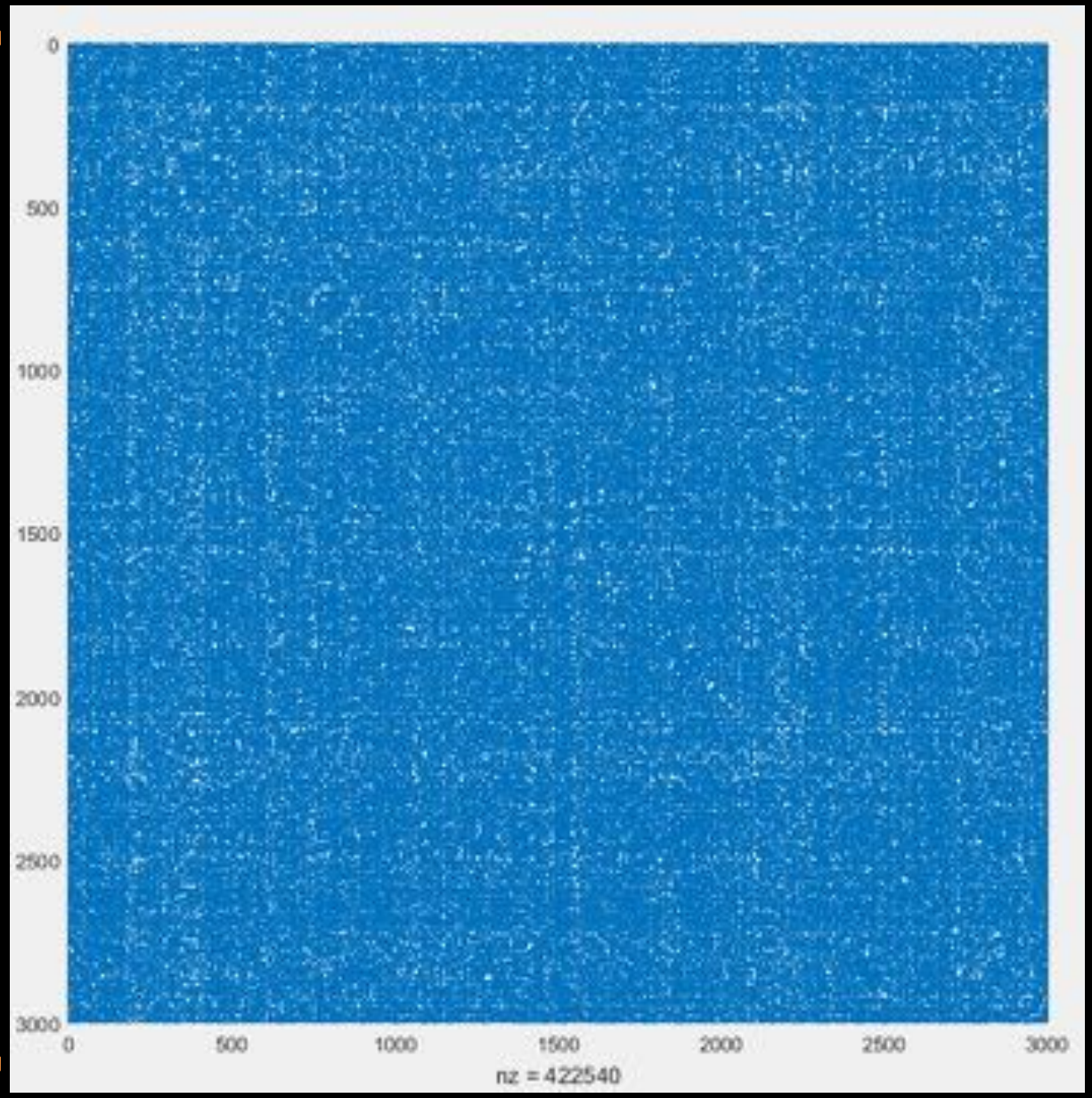


Graph Theory Based Unsupervised Learning

- Populate adjacency matrix

$$A_{ij} = \begin{cases} 1 & \text{if} \quad \Delta SD < \delta \\ 0 & \text{if} \quad \Delta SD > \delta \end{cases}$$

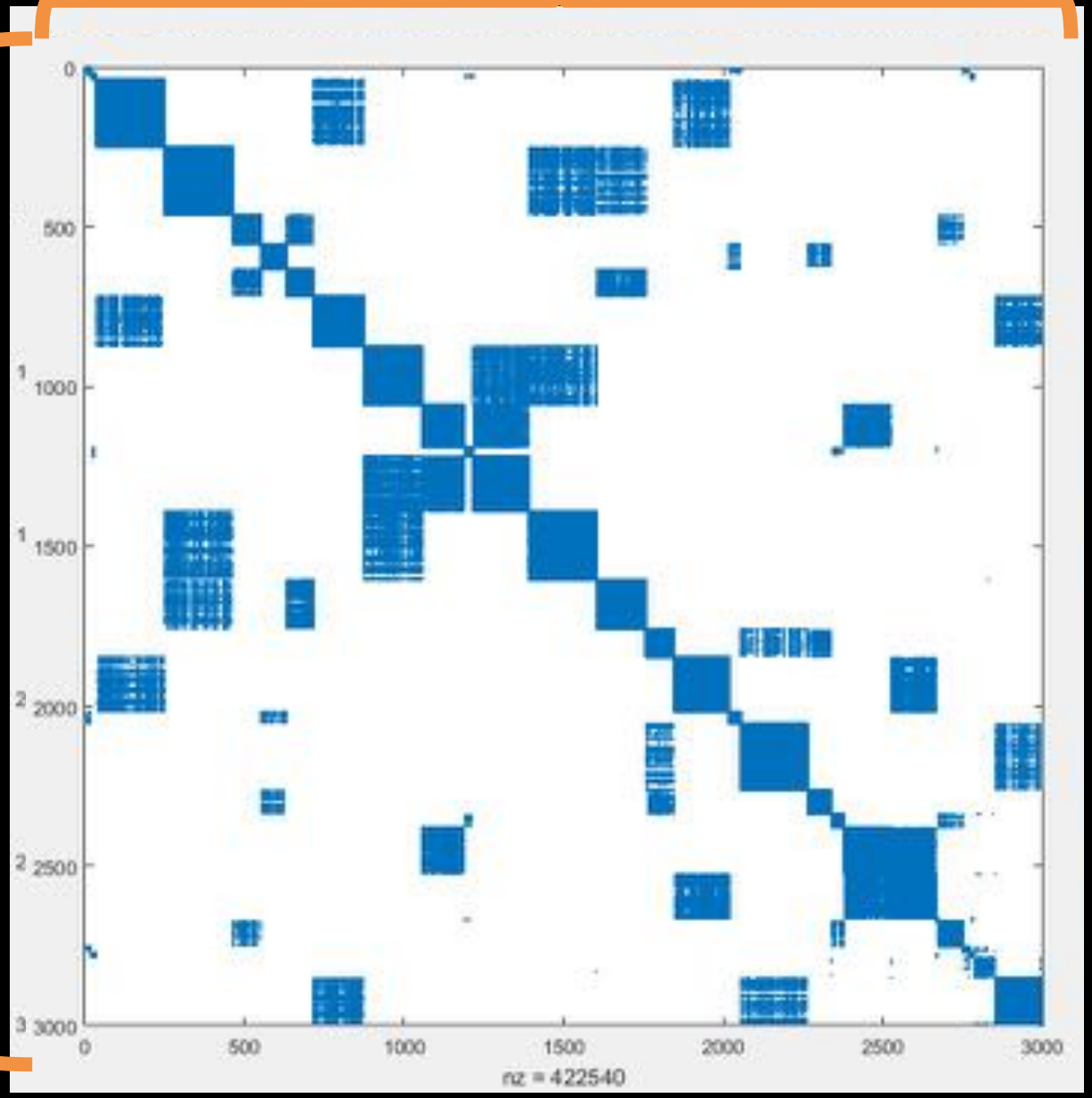
Nodes



Adjacency Matrix

Clusters

Nodes



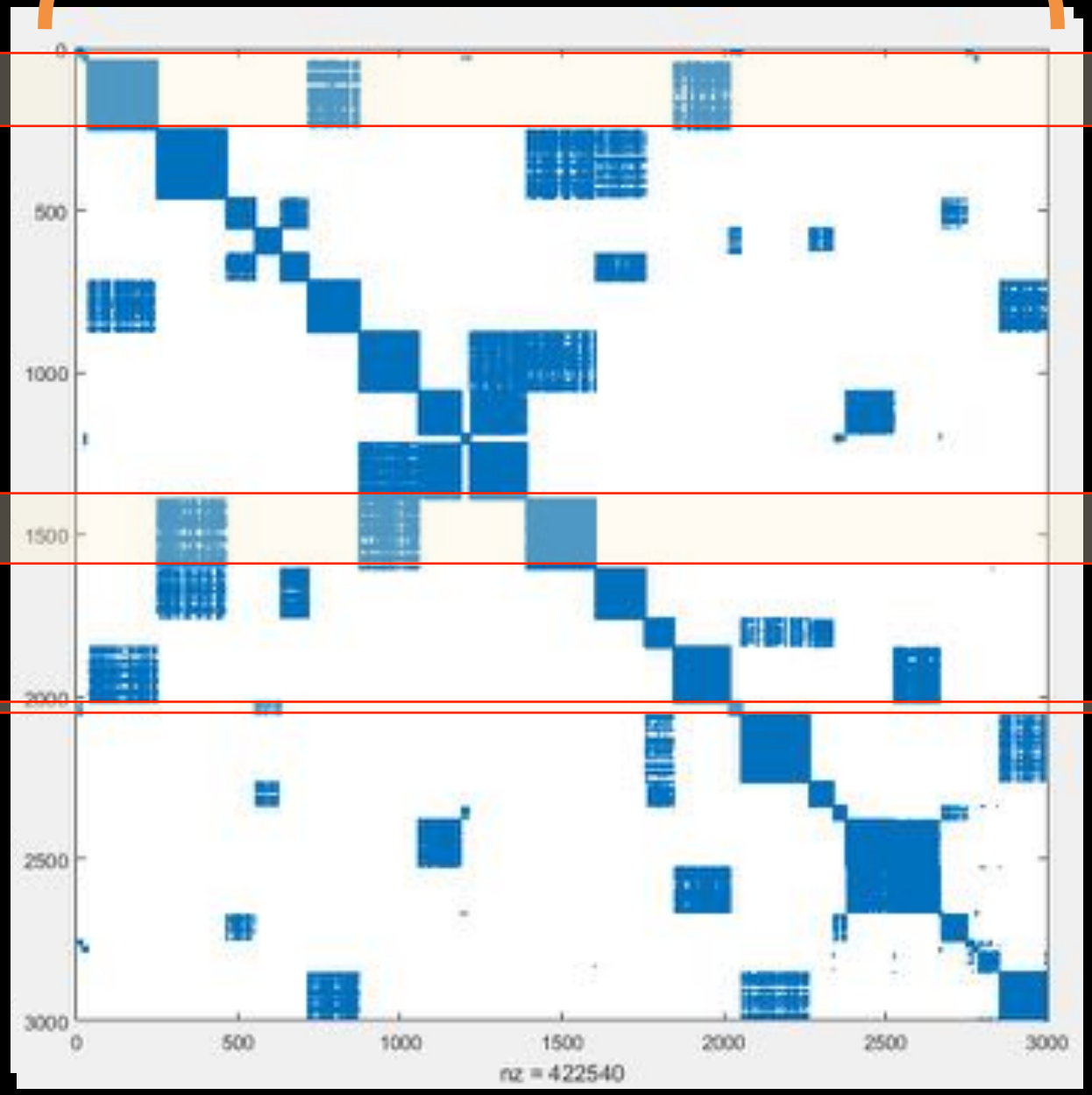
Adjacency Matrix

BOISE STATE UNIVERSITY

cryo GARS
CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

Clusters

CLUSTERS



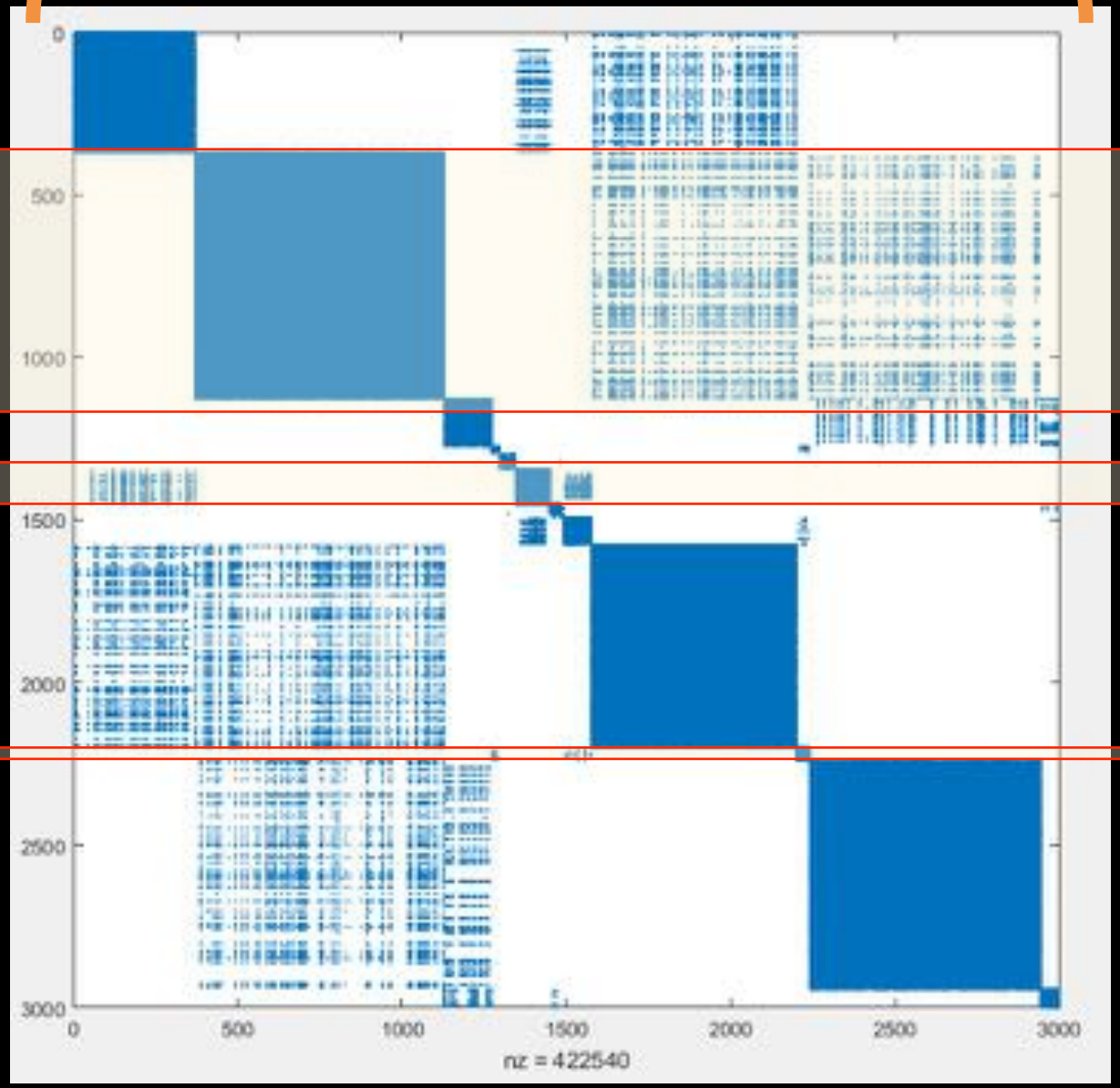
35 Total

BOISE STATE UNIVERSITY

cryo GARS
CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

Clusters

CLUSTERS



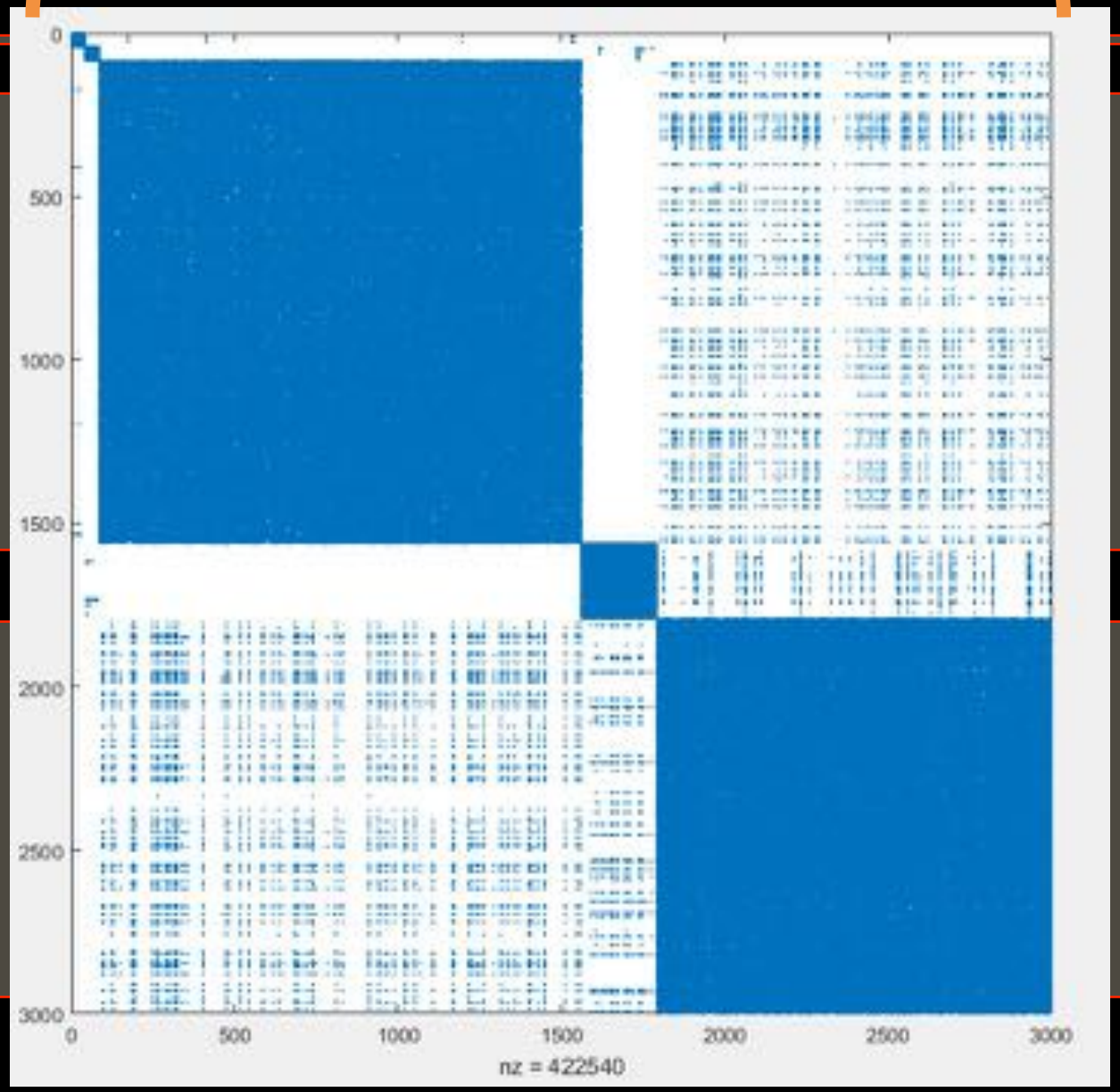
15 Total

BOISE STATE UNIVERSITY

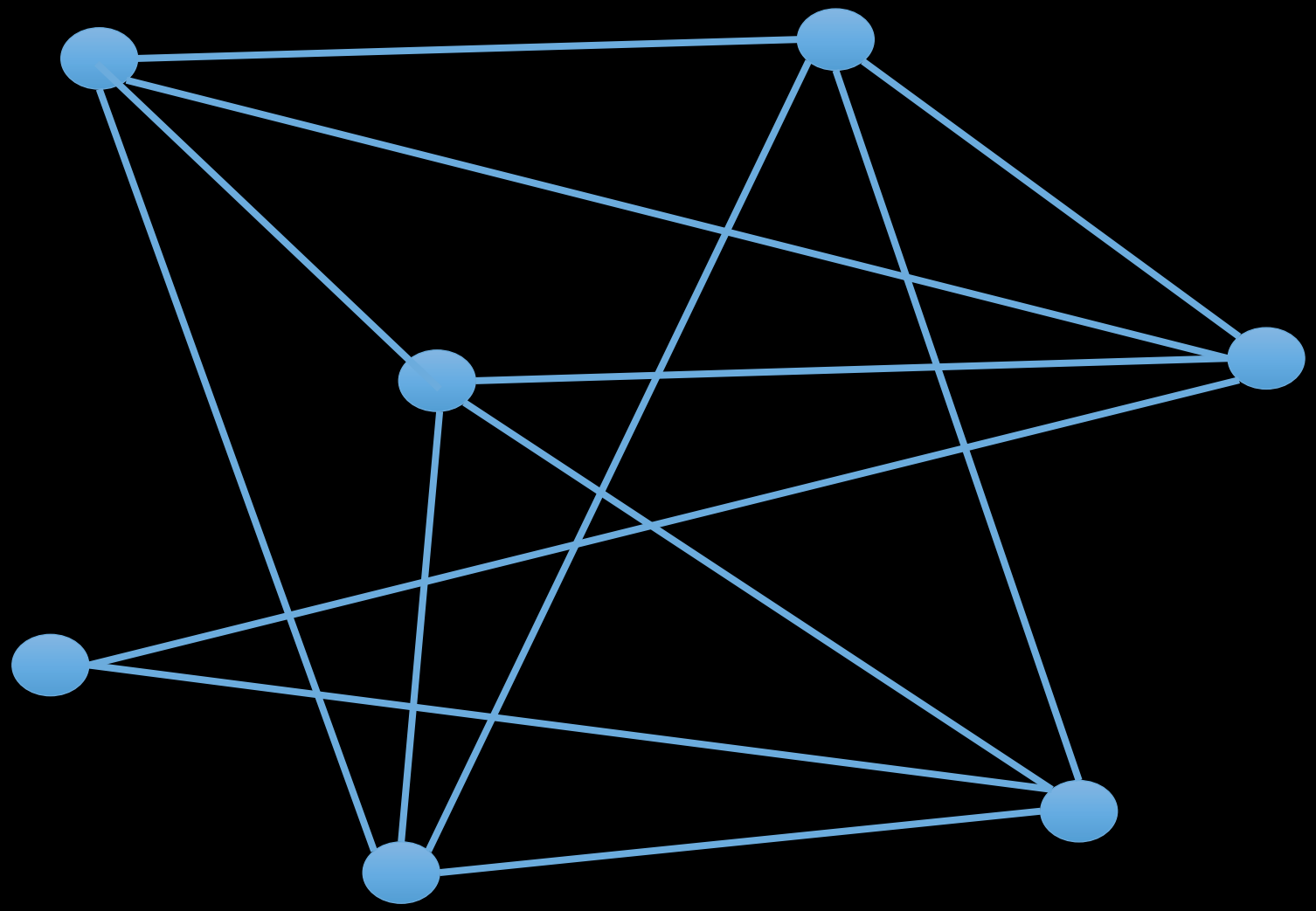
cryo GARS
CRYOSPHERE GEOPHYSICS AND REMOTE SENSING

Clusters

CLUSTERS

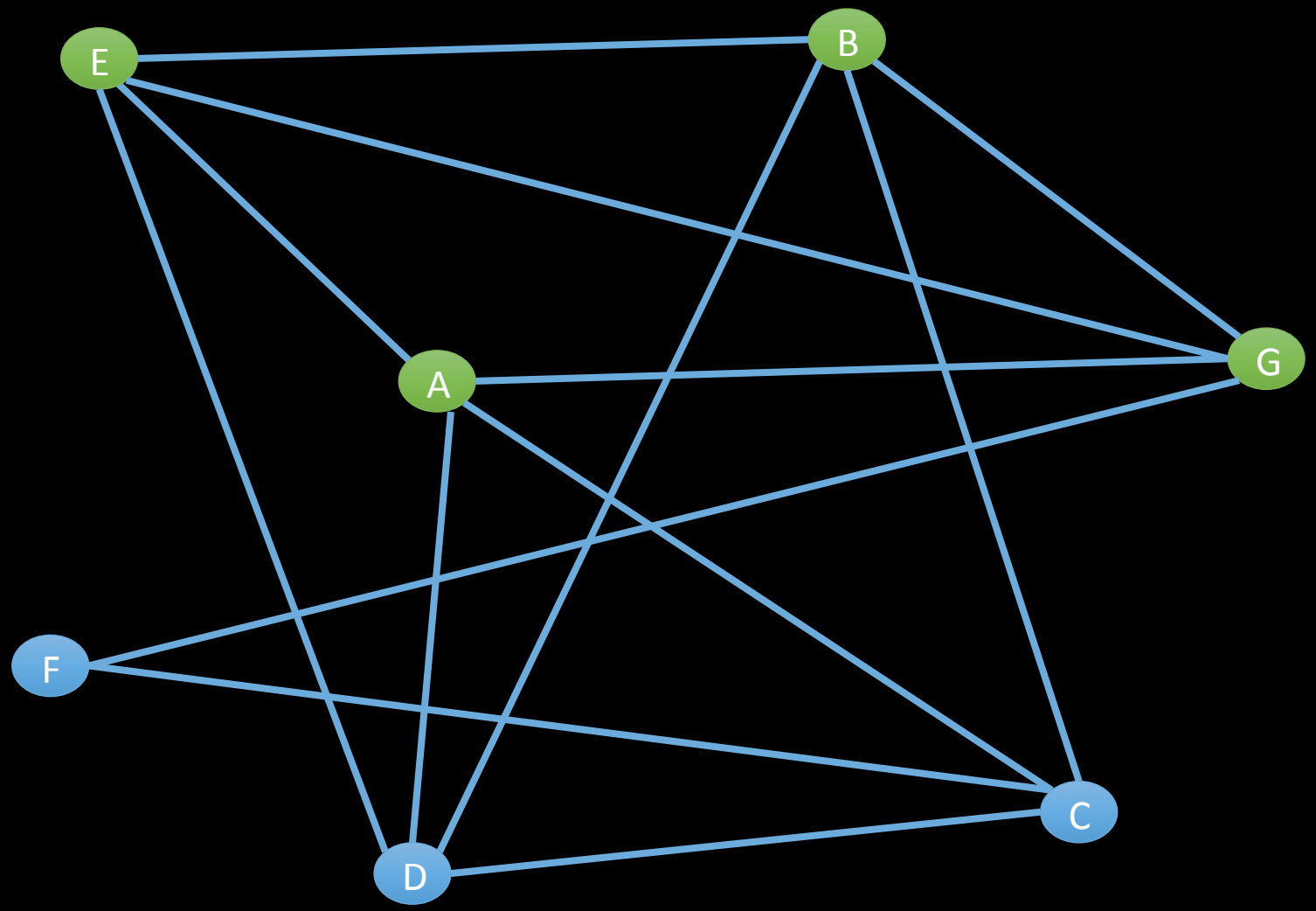


5 Total



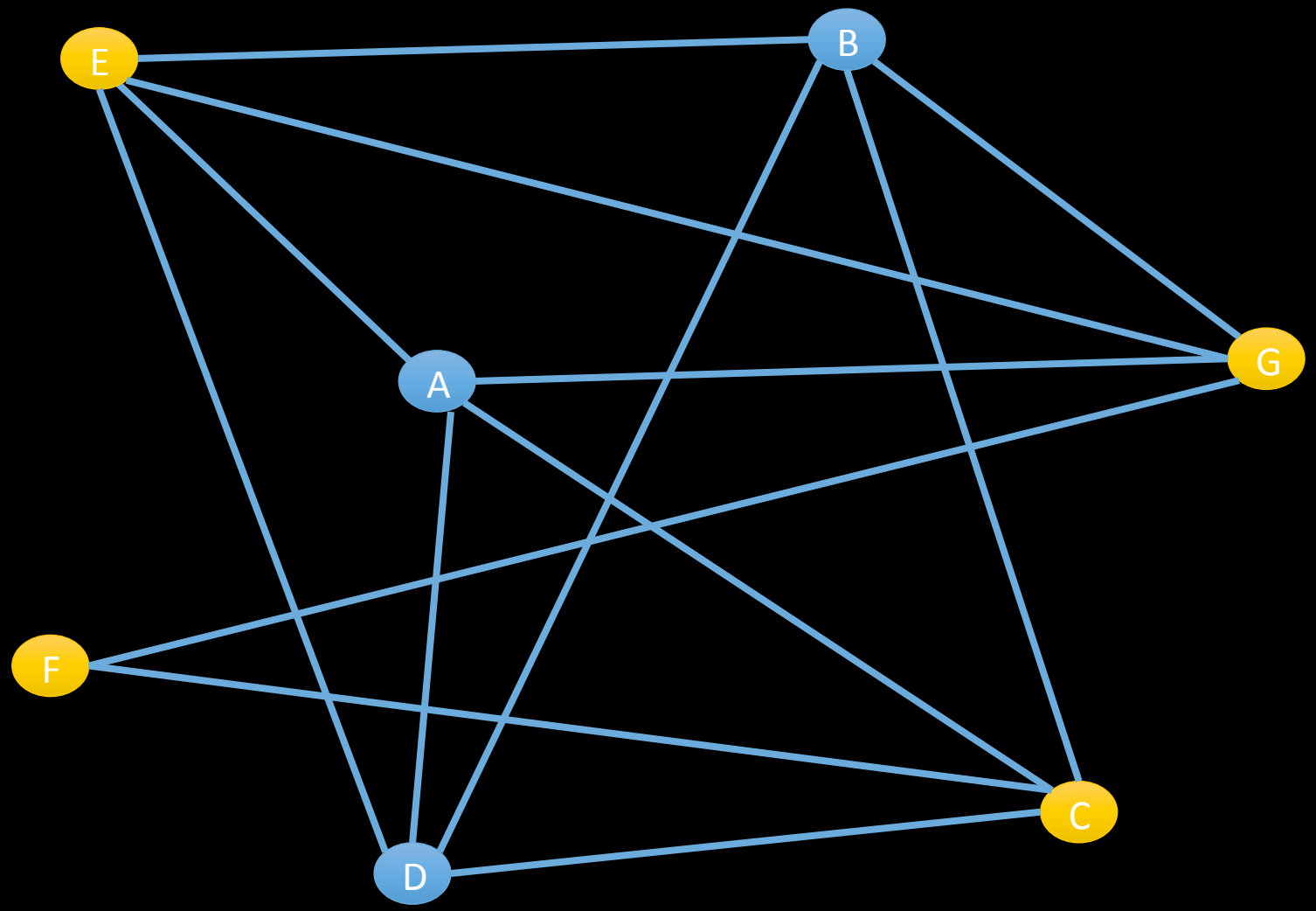
$$Q = \frac{1}{2m}\sum_{vw}\left[A_{vw} - \frac{k_v k_w}{2m}\right]\delta\left(c_v, c_w\right)$$

$$Q = 0.0118$$



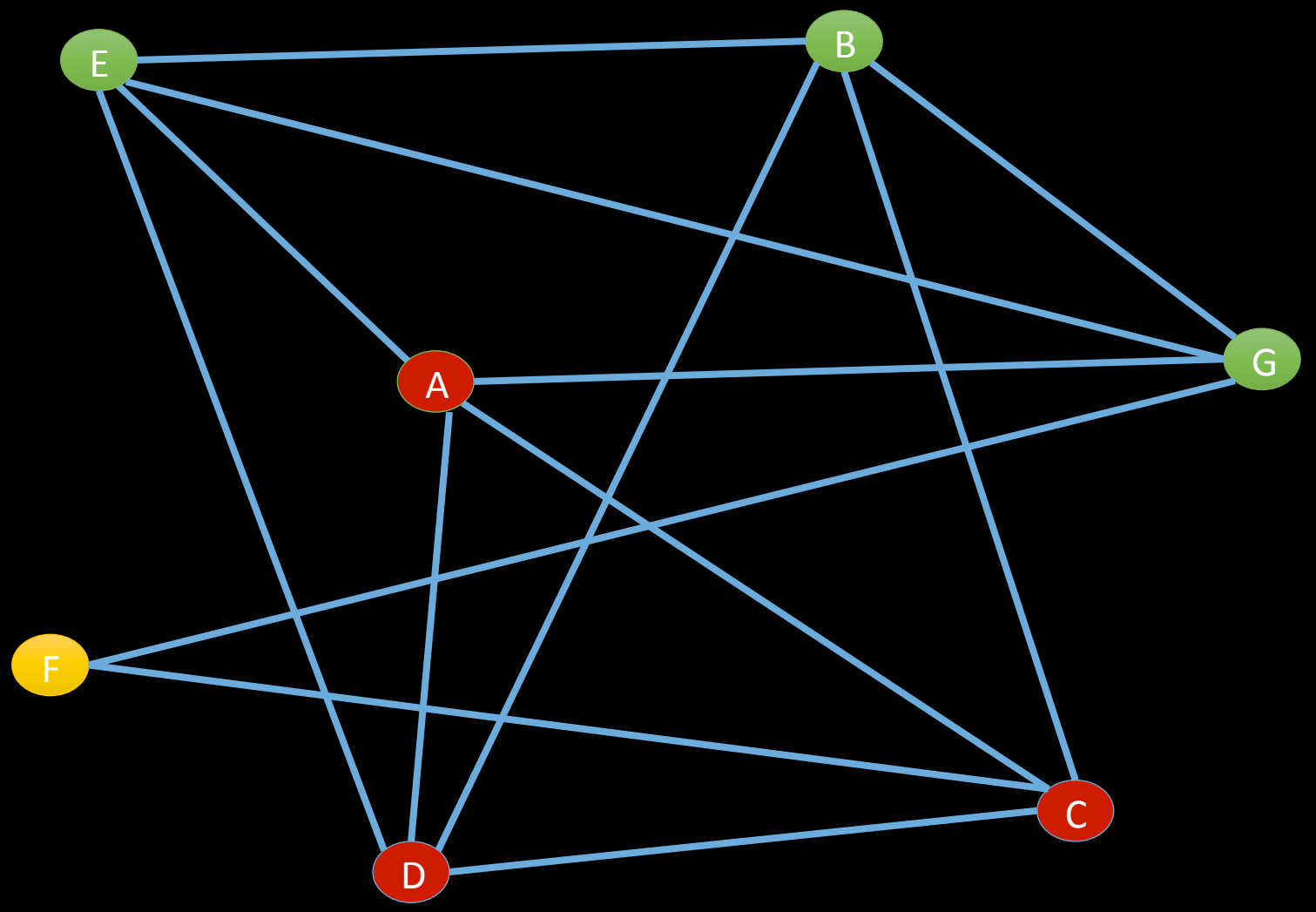
$$Q = \frac{1}{2m}\sum_{vw}\left[A_{vw} - \frac{k_v k_w}{2m}\right]\delta\left(c_v, c_w\right)$$

$$Q = -0.1183$$



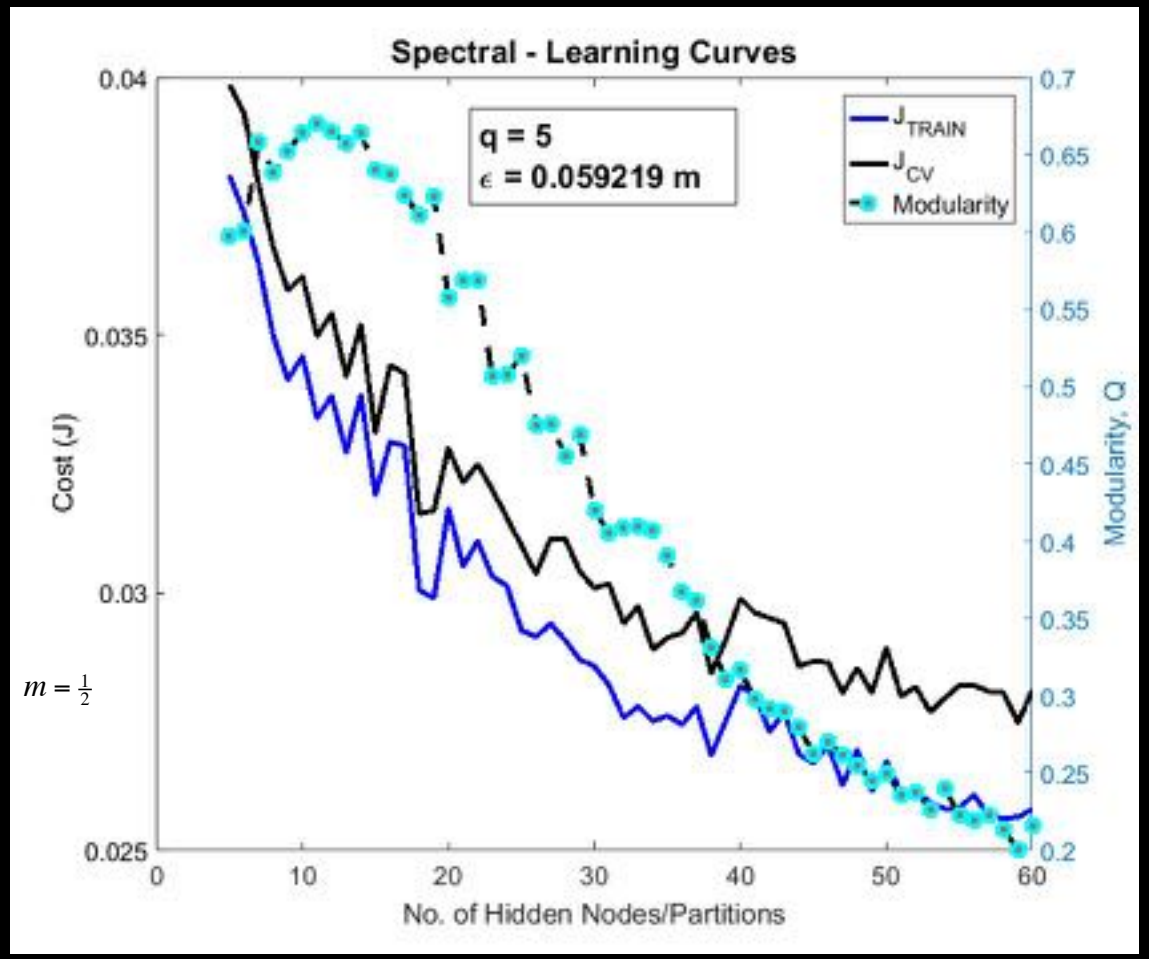
$$Q = \frac{1}{2m}\sum_{vw}\left[A_{vw} - \frac{k_v k_w}{2m}\right]\delta\left(c_v, c_w\right)$$

$$Q = 0.0296$$



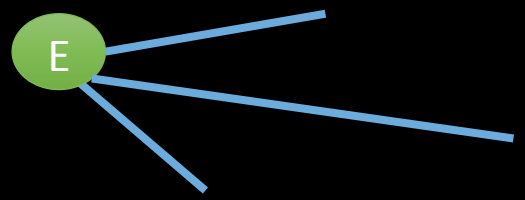
Use statistics from clusters to perform unsupervised training on the RBFNN.

Modularity



$m = \frac{1}{2}$

$$Q = \frac{1}{2m} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{2m} \right] \delta(c_v, c_w)$$



Clauset et al 2004

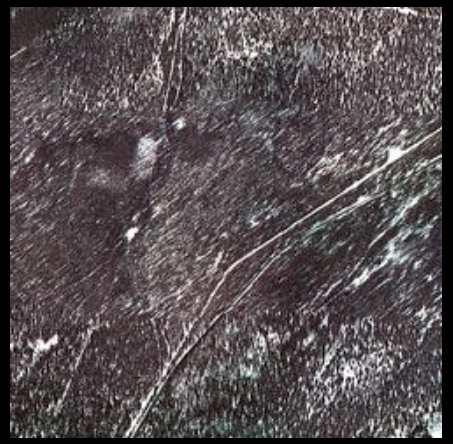
Data

CLPX-I

Spring Creek

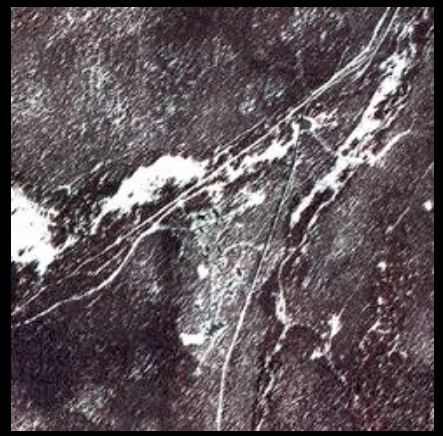
Buffalo Pass

Walton Creek



St. Louis Creek

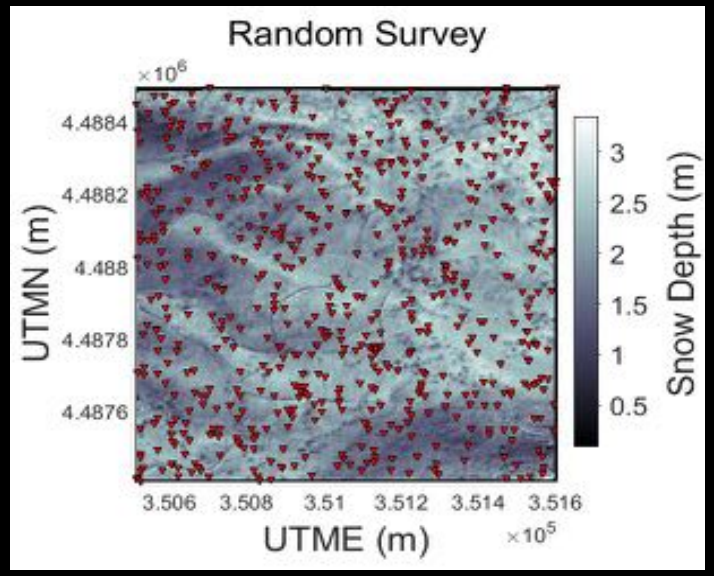
Alpine

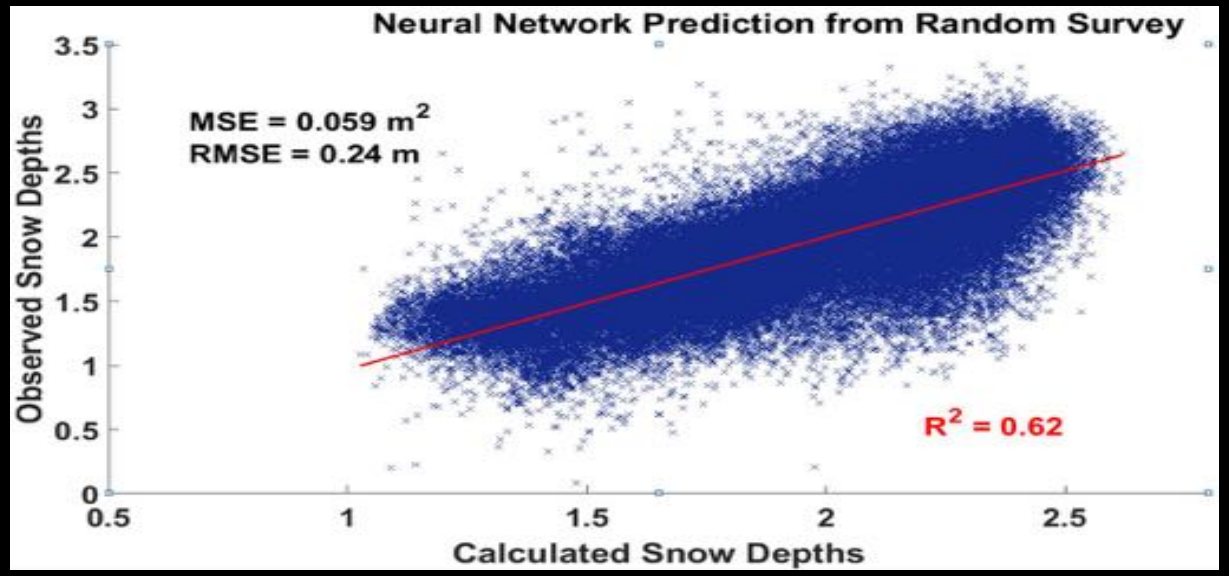


LSOS

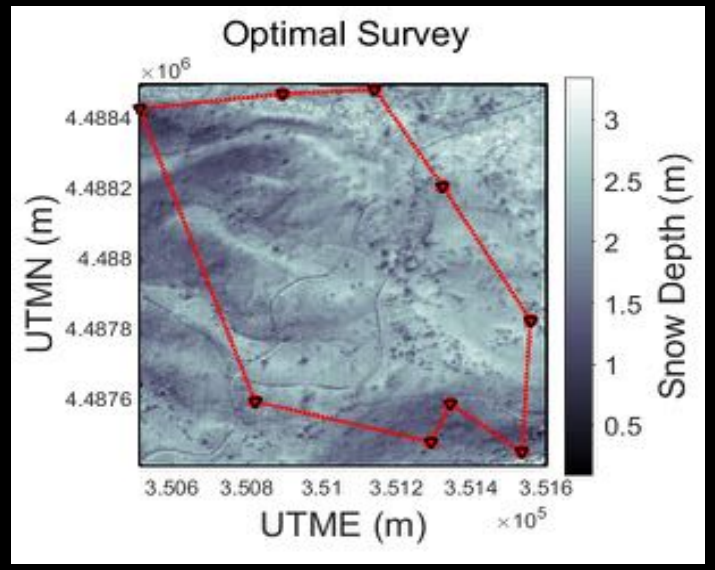
Preliminary Results

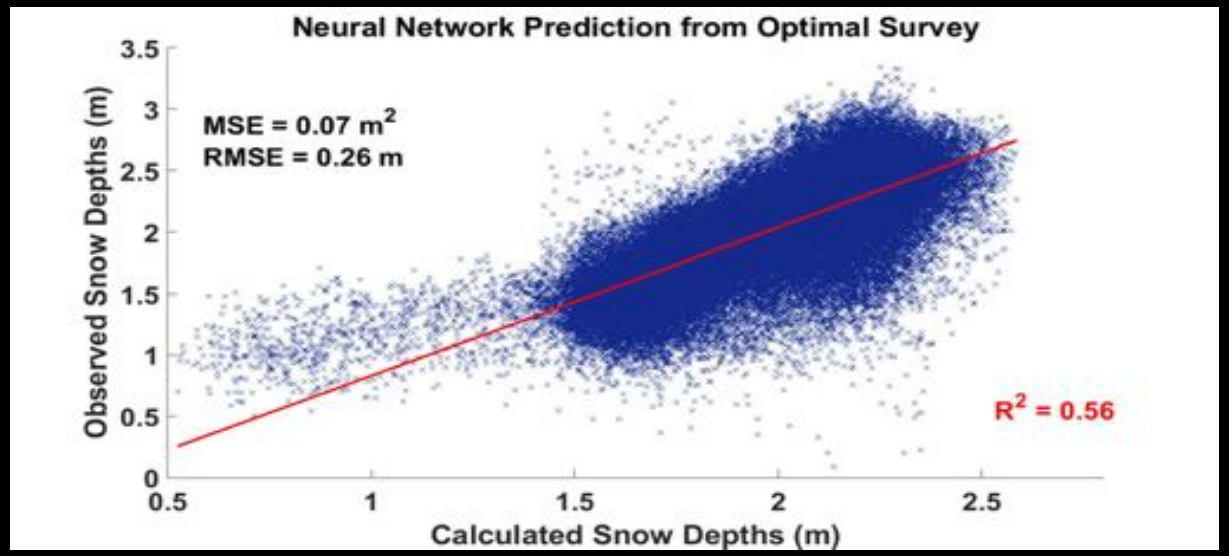
600 Random Points

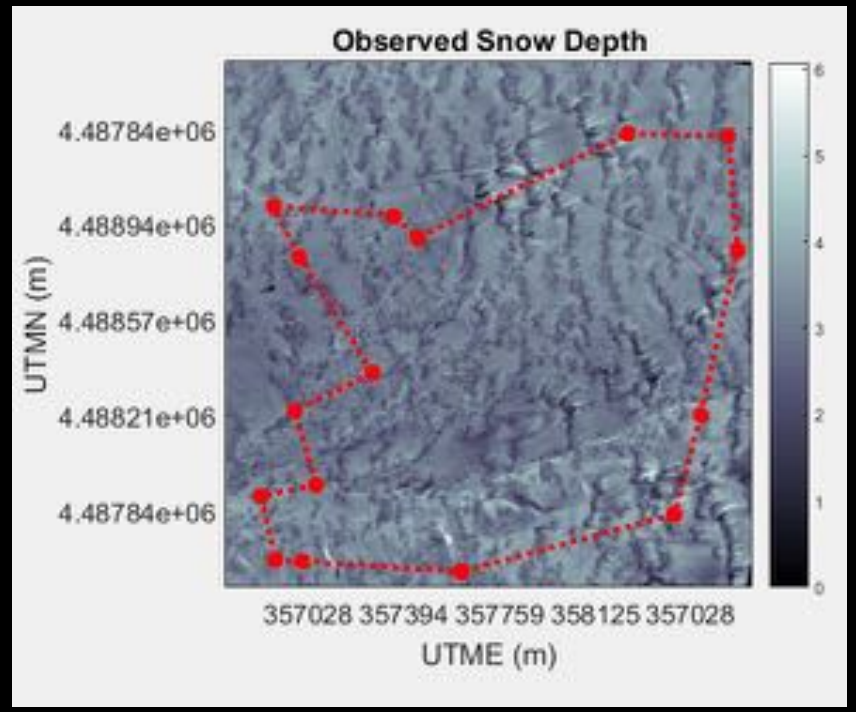


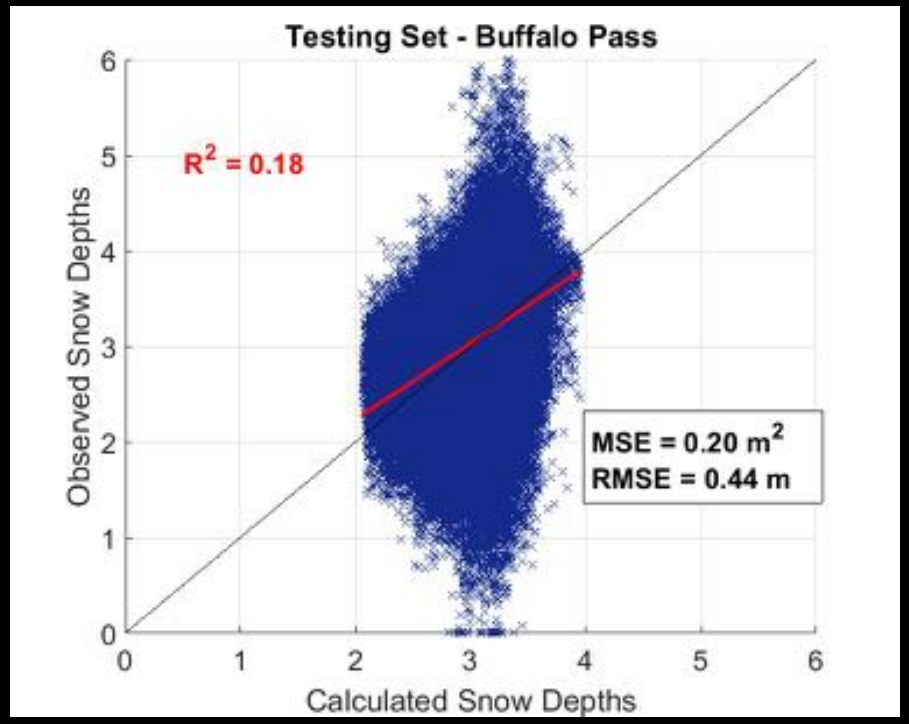


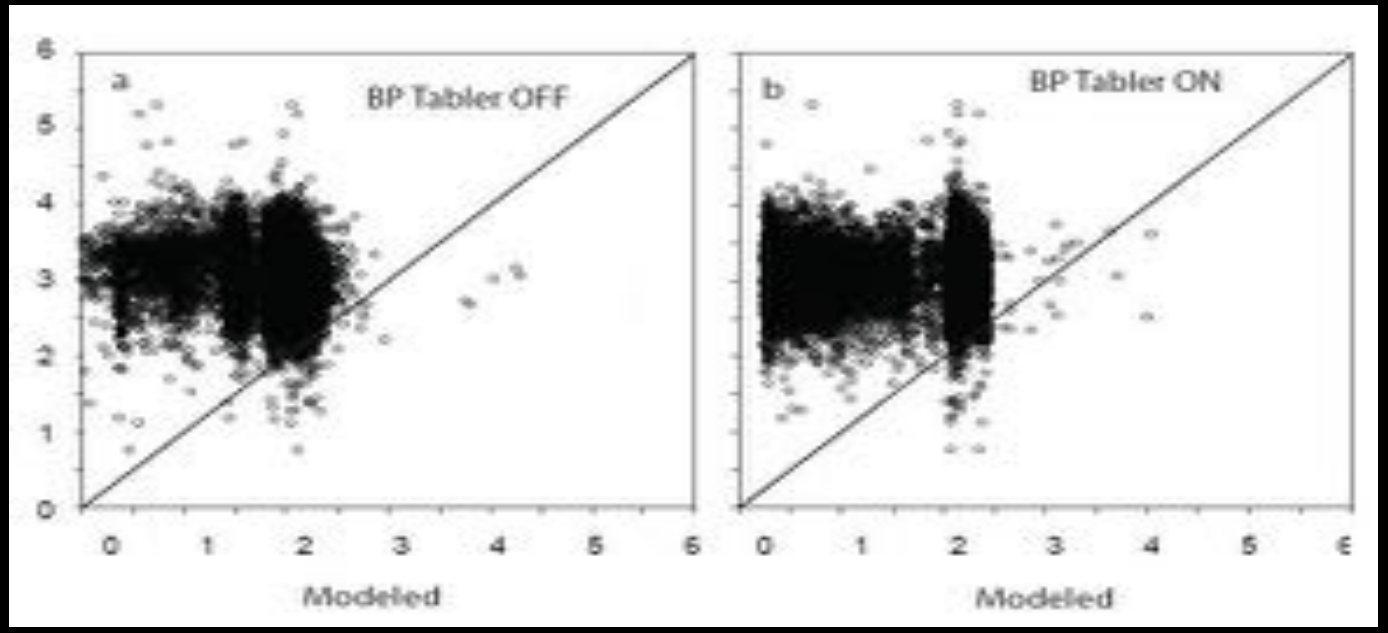
9 High Betweeness Points





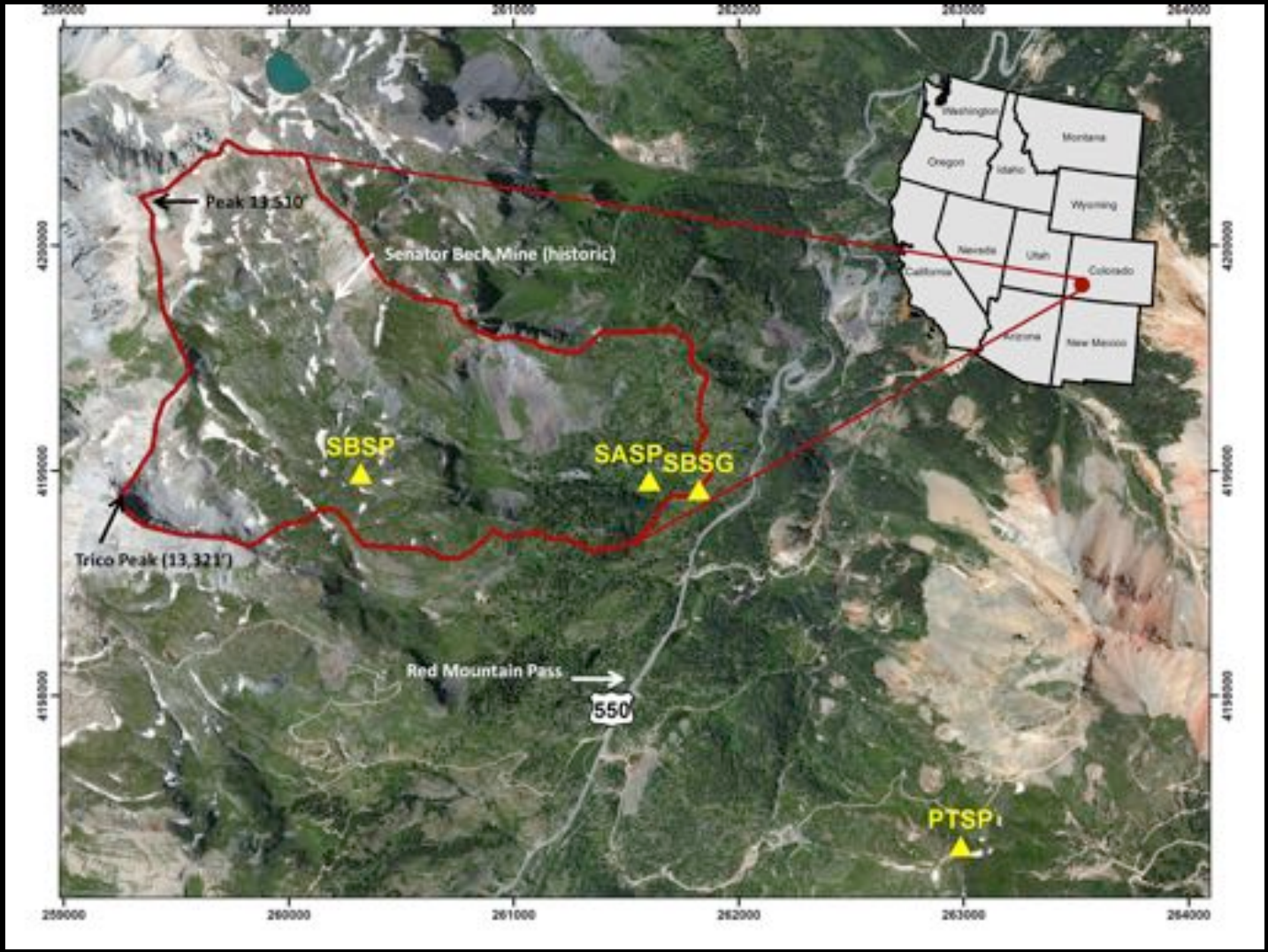

Observed Snow Depth
4.48784e+06
4.48894e+06
4.48857e+06
4.48821e+06
4.48784e+06
UTMN (m)
357028 357394 357759 358125 357028
UTME (m)


Testing Set - Buffalo Pass
R^2 = 0.18
MSE = 0.20 m^2
RMSE = 0.44 m
Observed Snow Depths
Calculated Snow Depths


a
BP Tabler OFF
b
BP Tabler ON
Modeled
Modeled

Deems 2009

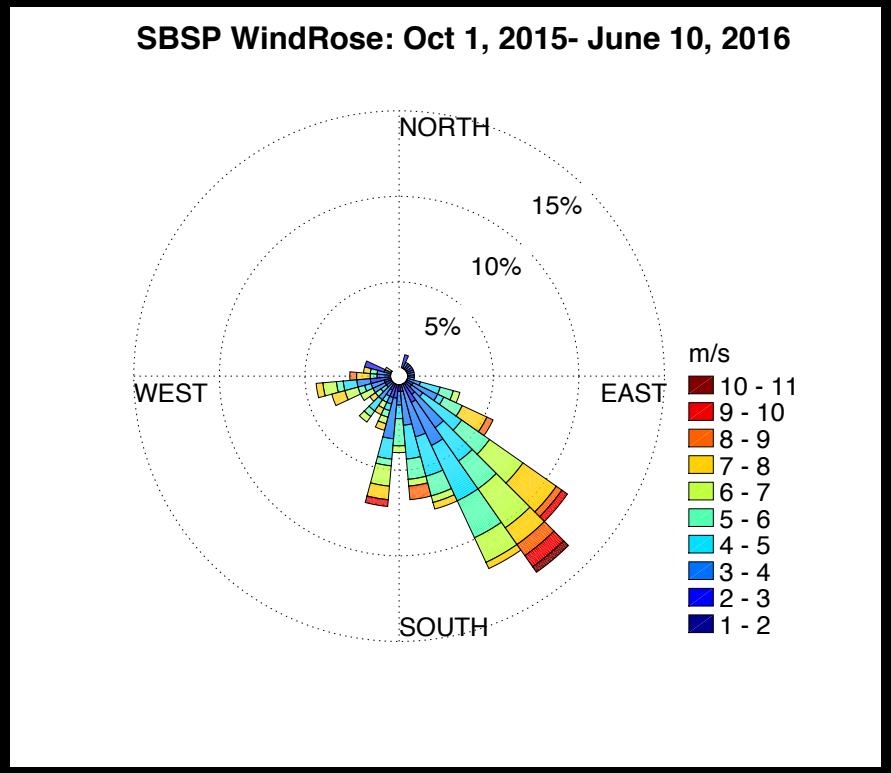
Senator Beck Basin CO

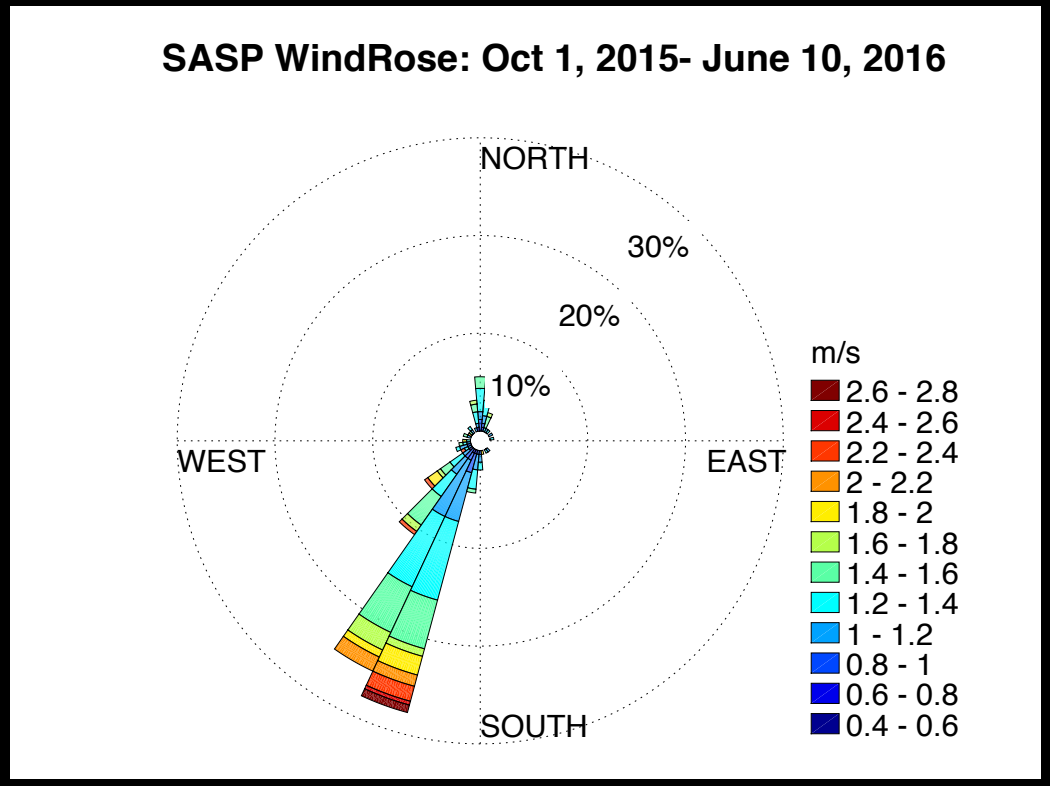


Center for Snow & Avalanche Studies

Wind Direction

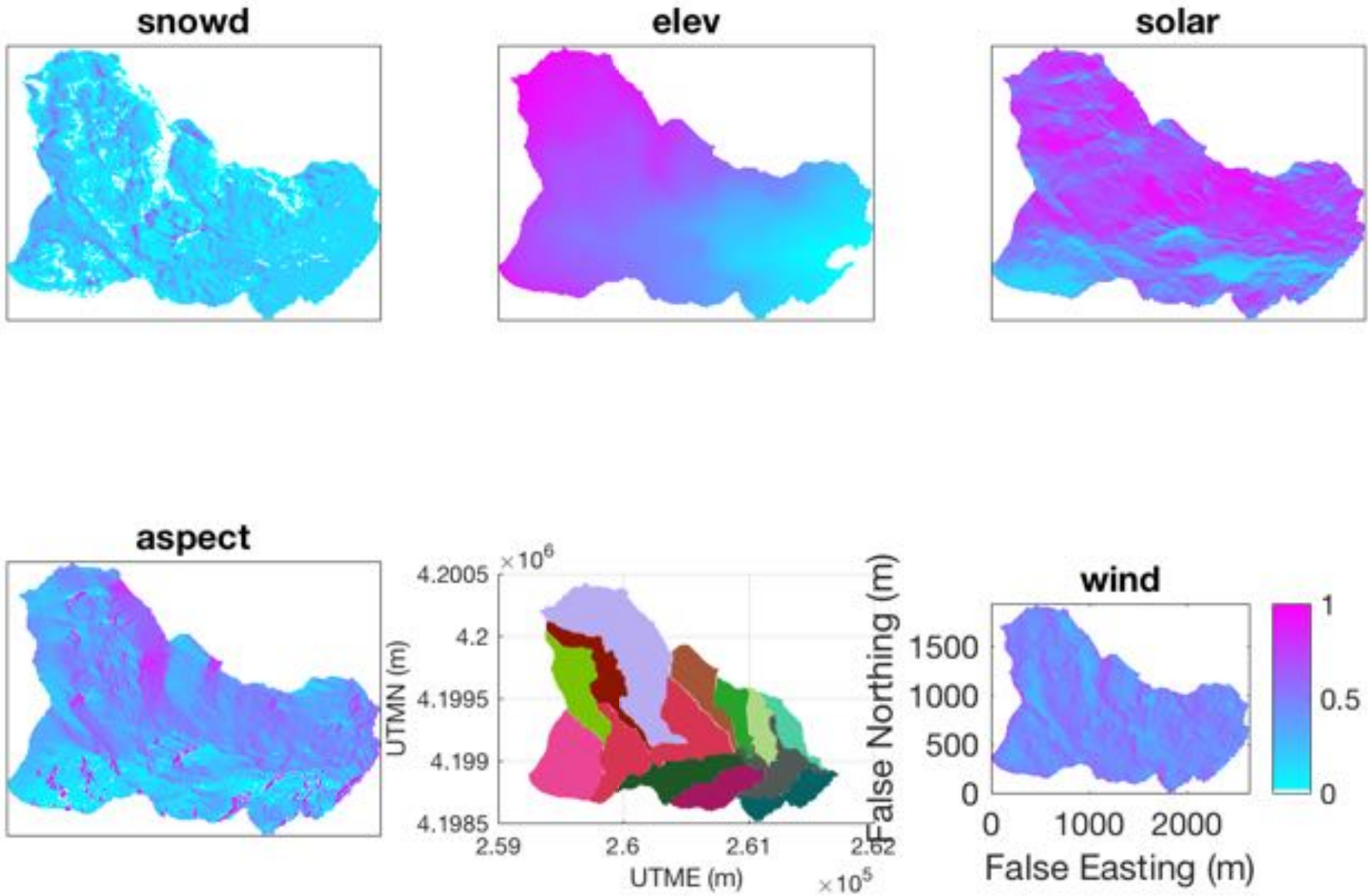
(Winstral et. al 2002)

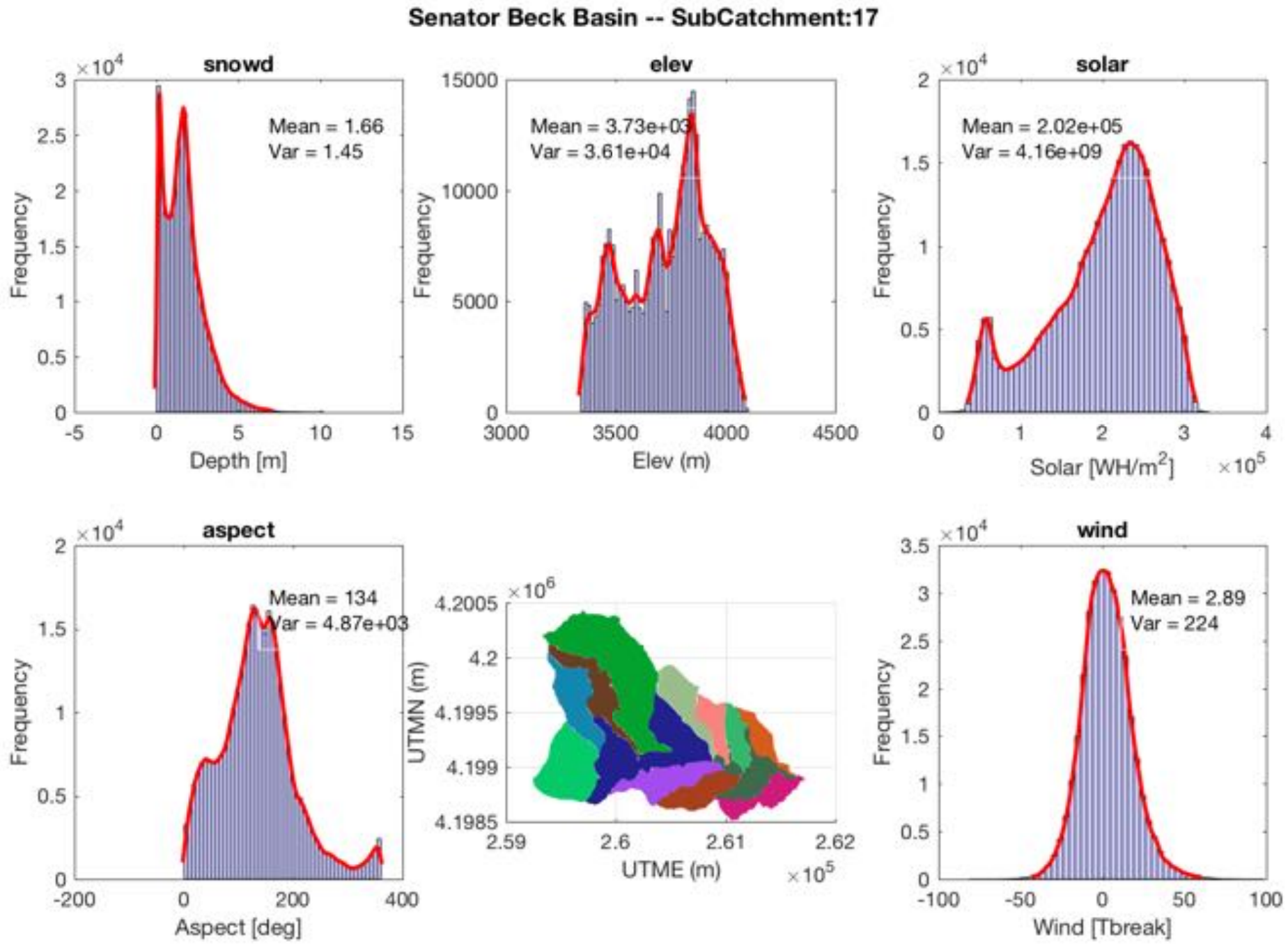




Center for Snow & Avalanche Studies

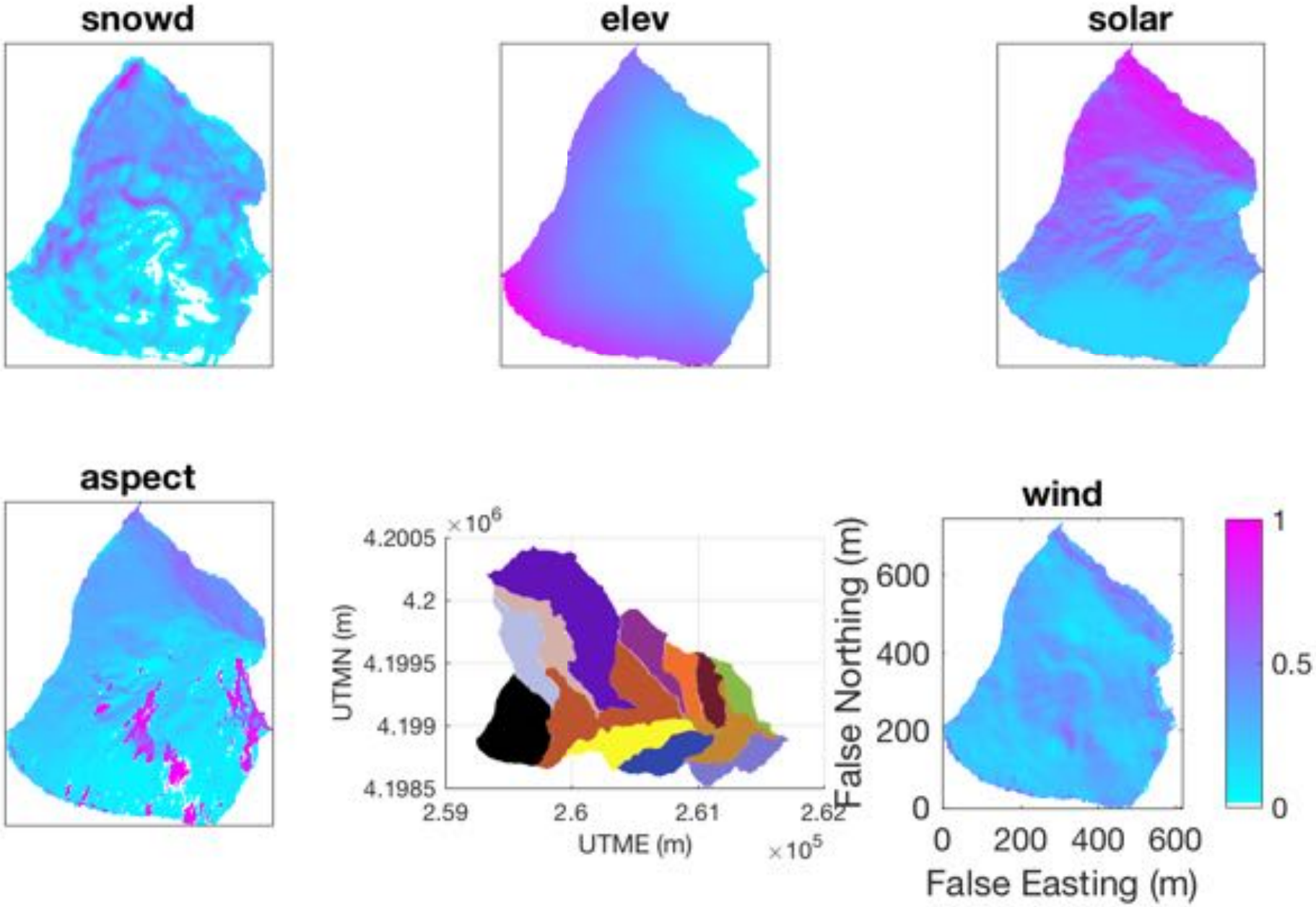
Senator Beck Basin -- SubCatchment:17

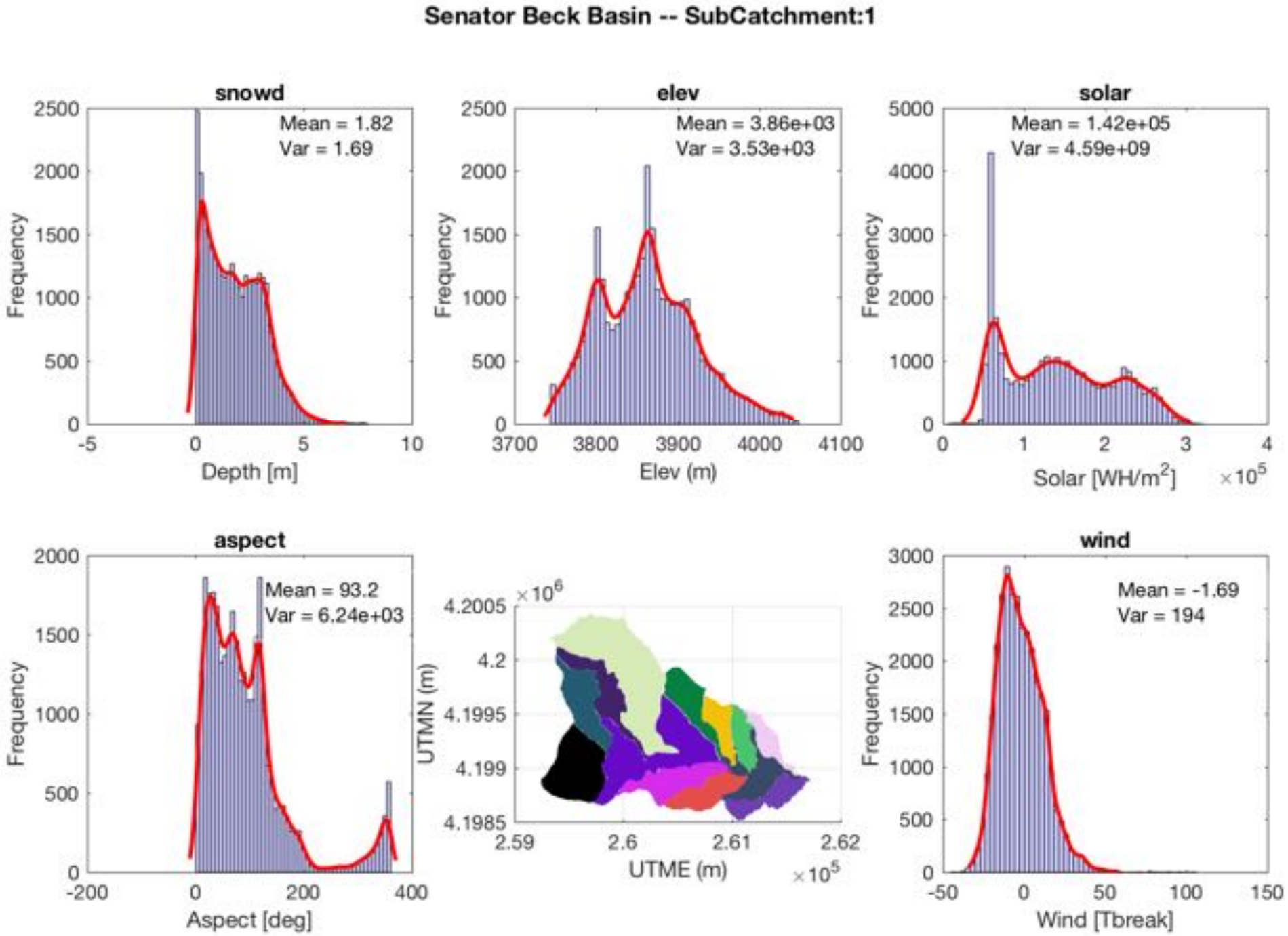



Senator Beck Basin -- SubCatchment:17
snowd
Mean = 1.66
Var = 1.45
Frequency
Depth [m]
elev
Mean = 3.73e+03
Var = 3.61e+04
Frequency
Elev (m)
solar
Mean = 2.02e+05
Var = 4.16e+09
Frequency
Solar [WH/m2]
aspect
Mean = 134
Var = 4.87e+03
Frequency
Aspect [deg]
UTMN (m)
UTME (m)
wind
Mean = 2.89
Var = 224
Frequency
Wind [Tbreak]

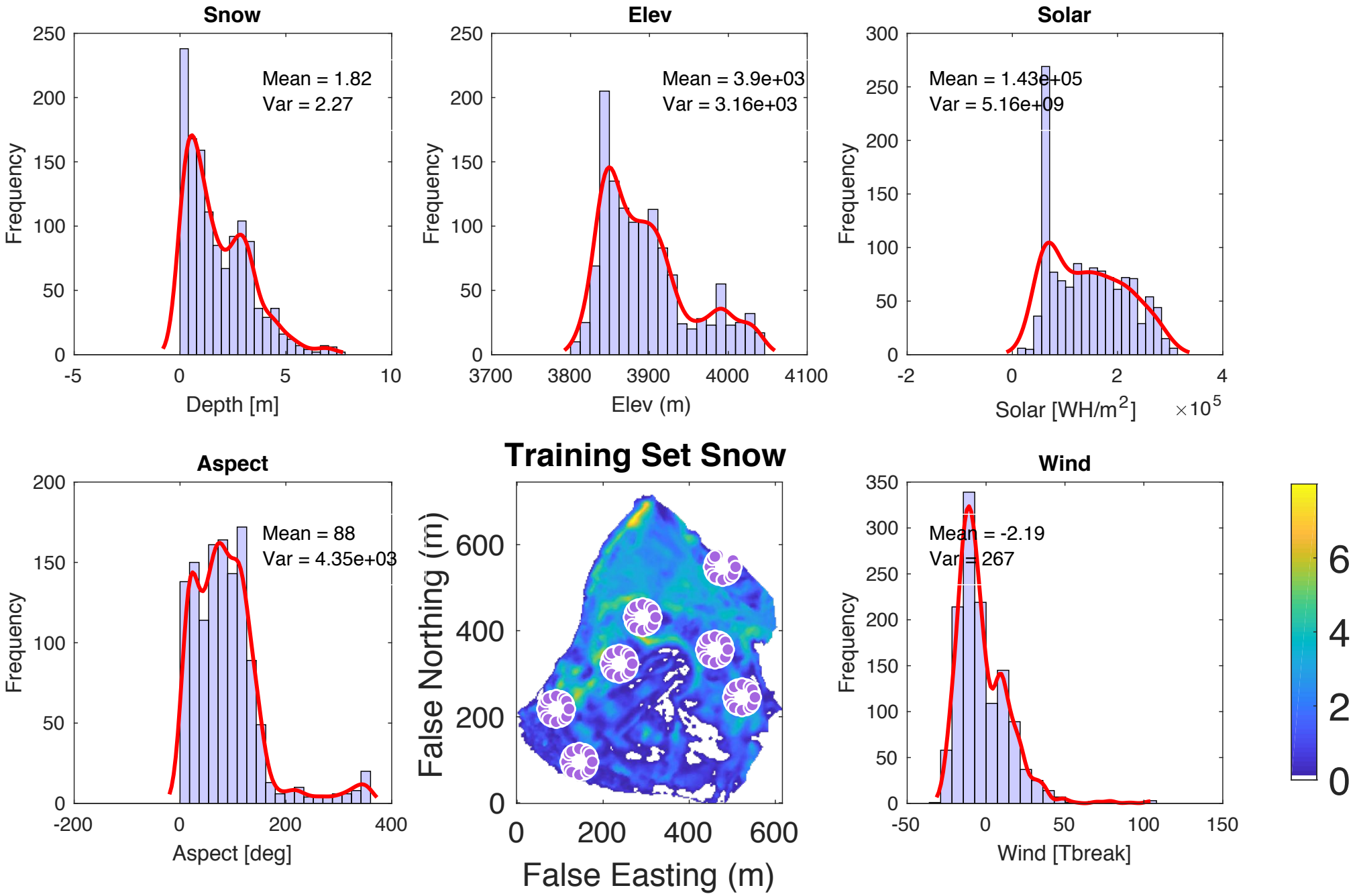
Sample Design

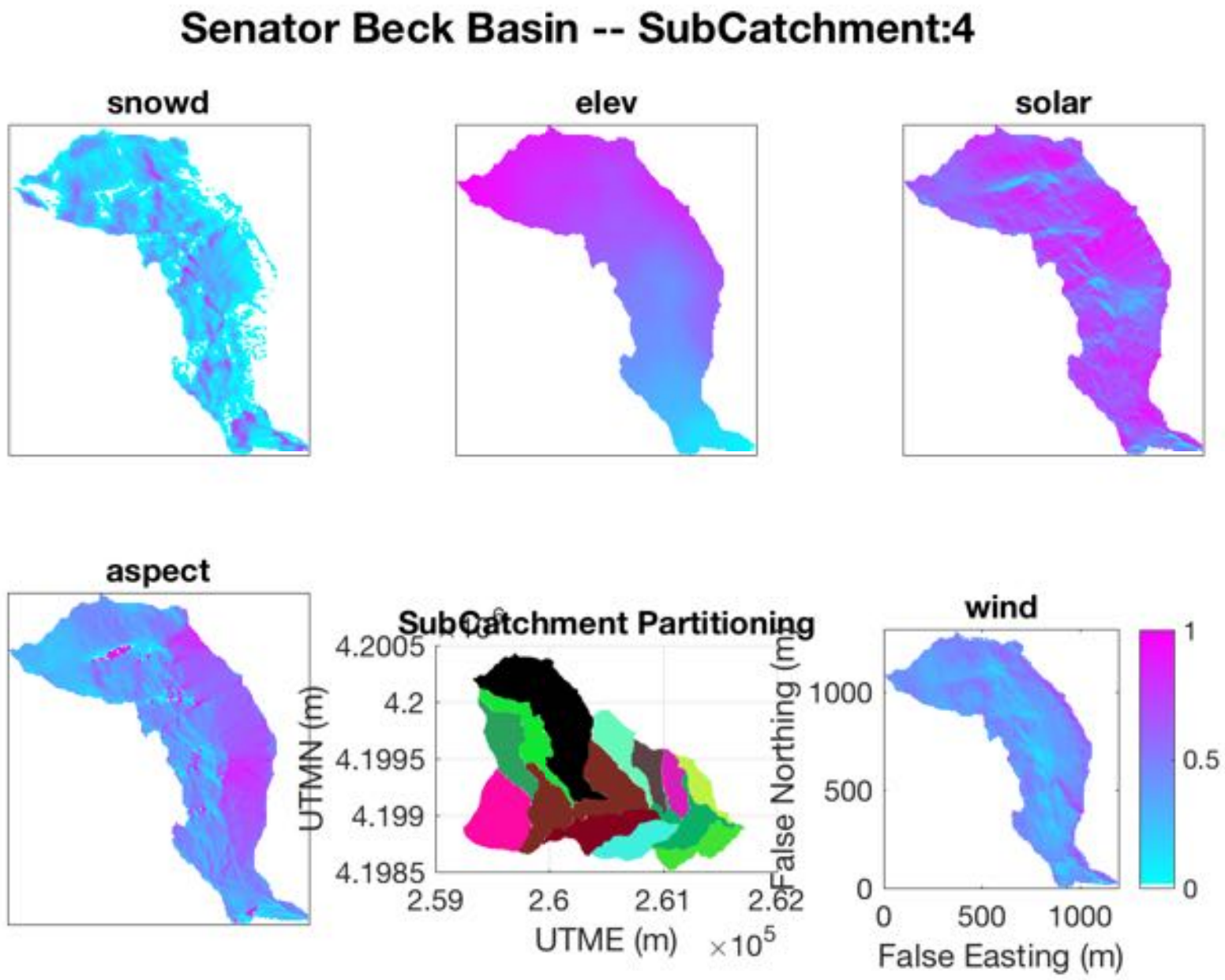
Senator Beck Basin -- SubCatchment:1

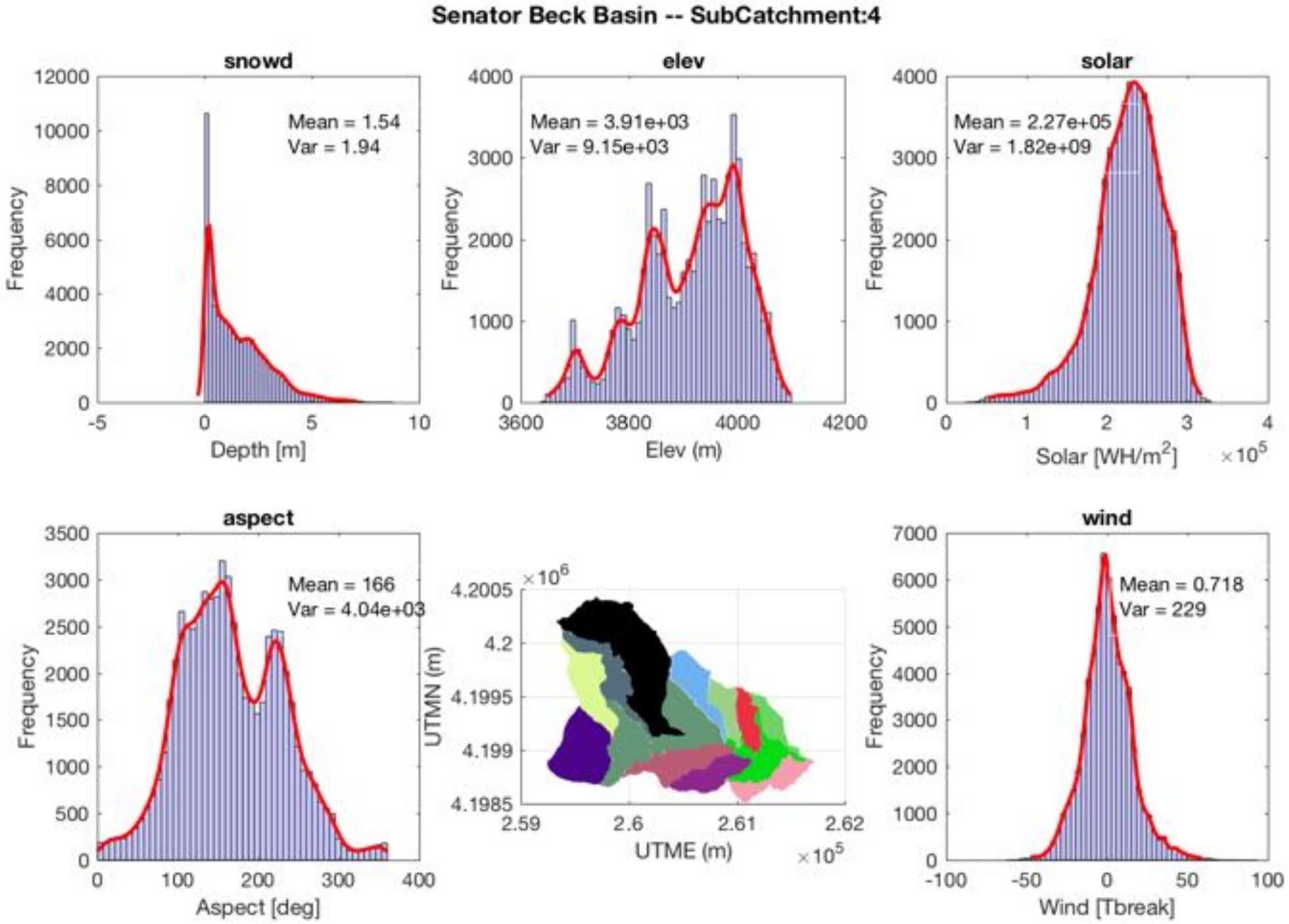



Senator Beck Basin -- SubCatchment:1
snowd
Mean = 1.82
Var = 1.69
Frequency
Depth [m]
elev
Mean = 3.86e+03
Var = 3.53e+03
Elev (m)
solar
Mean = 1.42e+05
Var = 4.59e+09
Solar [WH/m2]
×10^5
aspect
Mean = 93.2
Var = 6.24e+03
Aspect [deg]
×10^6
UTMN (m)
UTME (m)
×10^5
wind
Mean = -1.69
Var = 194
Wind [Tbreak]

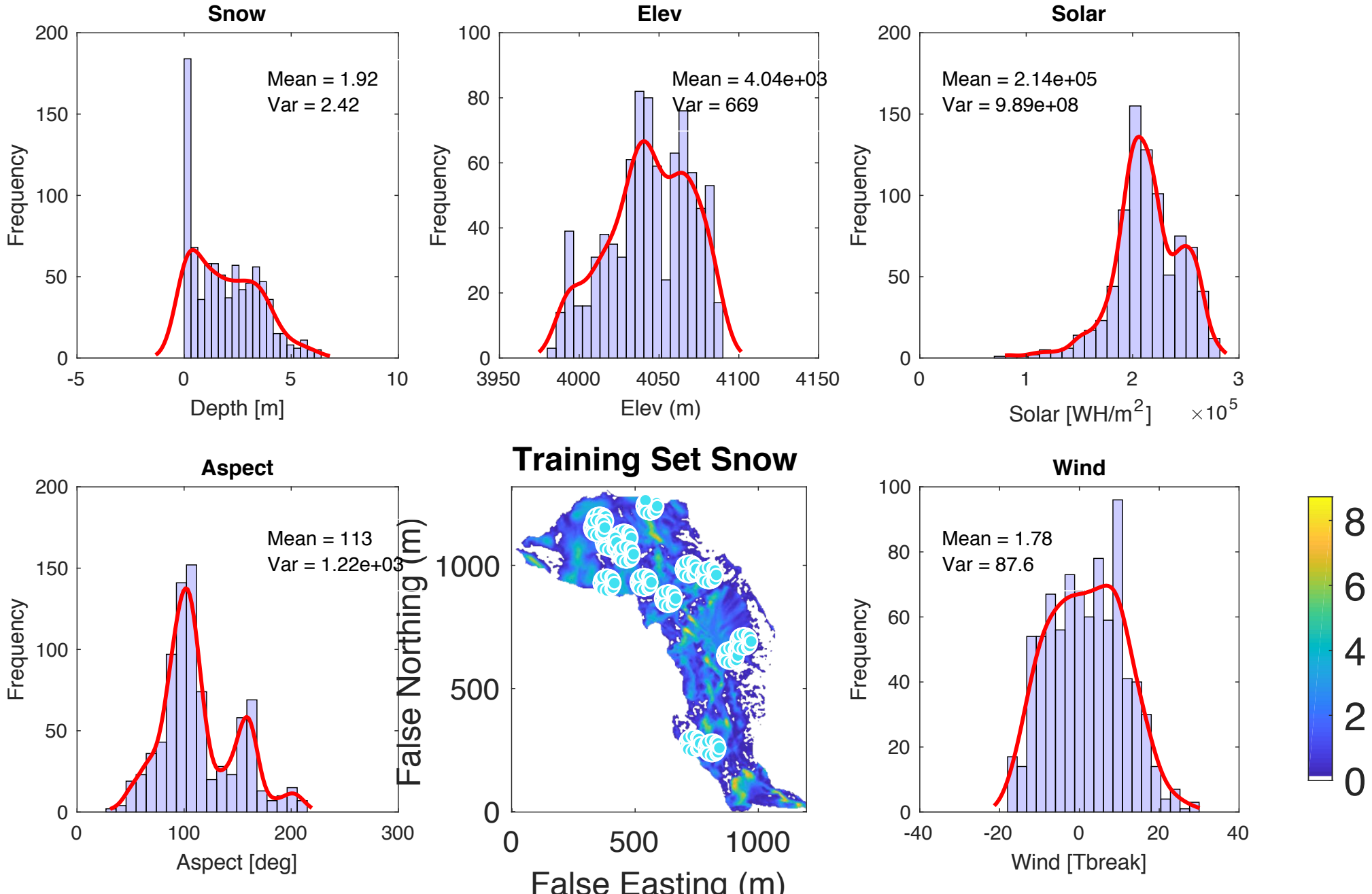
Senator Beck Basin -- SubCatchment:1



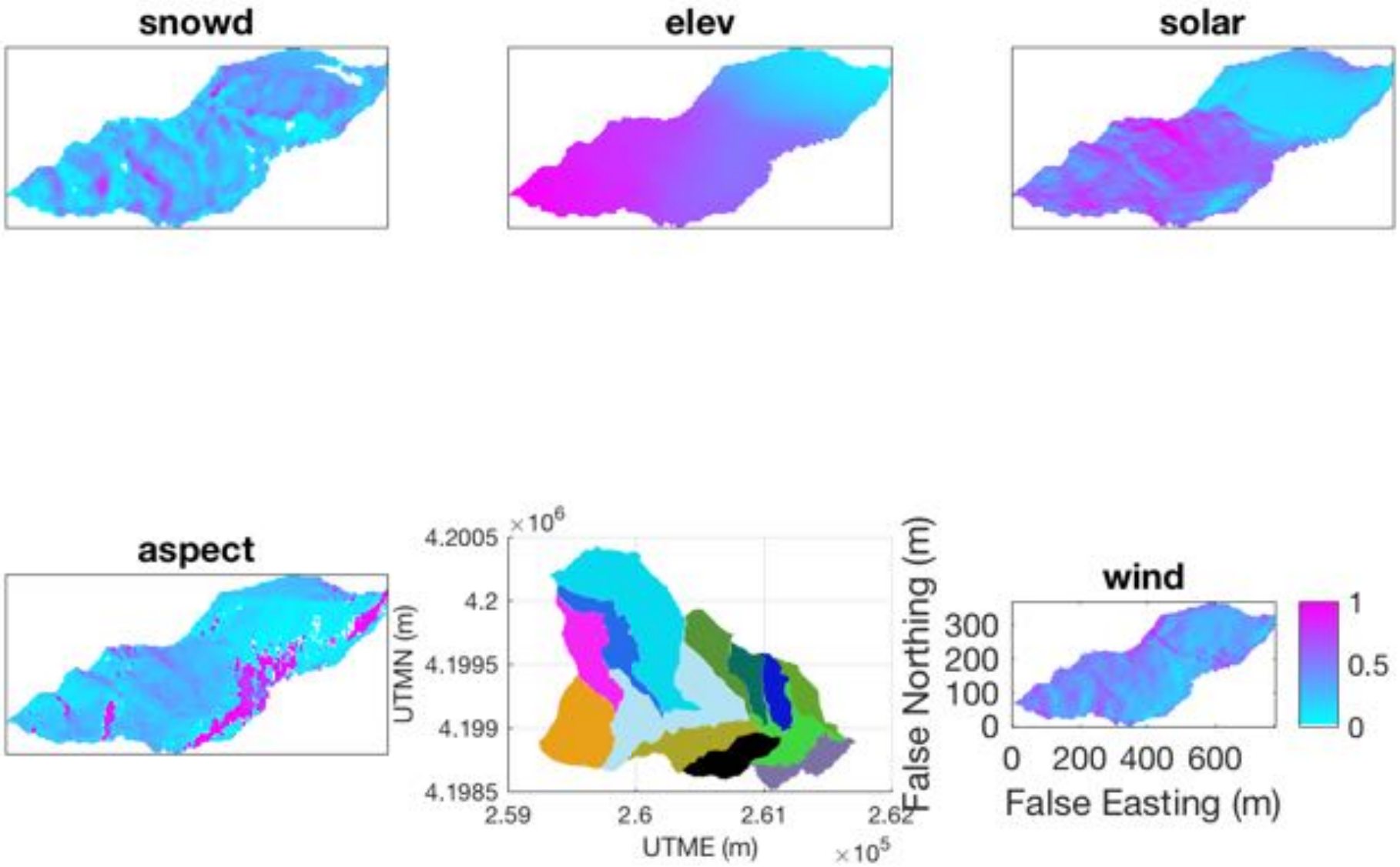

Senator Beck Basin -- SubCatchment:4
snowd
elev
solar
aspect
SubCatchment Partitioning
UTMN (m)
4.2005
4.2
4.1995
4.199
4.1985
2.59
2.6
2.61
2.62
UTME (m)
×10^5
wind
False Northing (m)
1000
500
0
0
500
1000
False Easting (m)
1
0.5
0

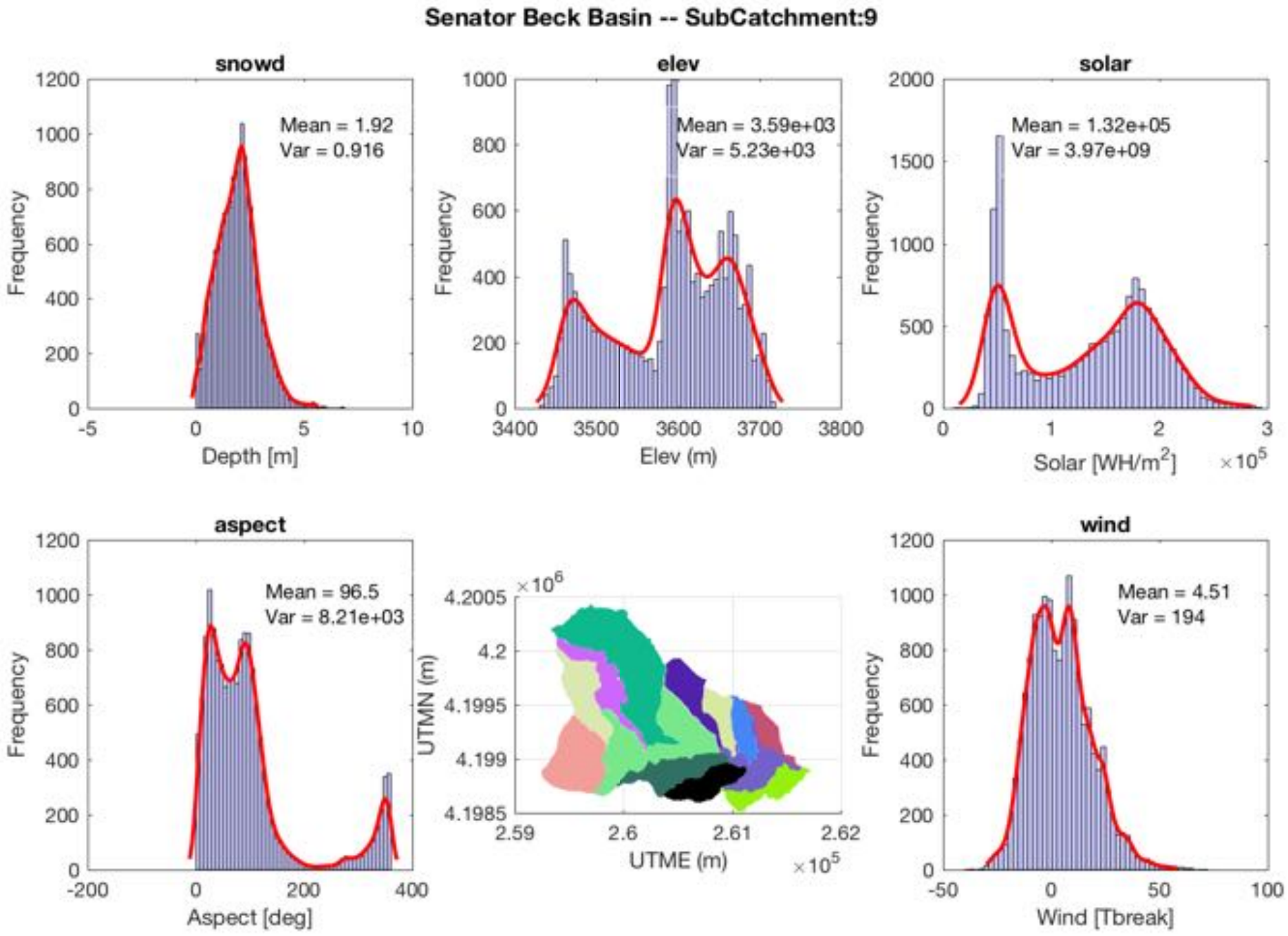

Senator Beck Basin -- SubCatchment:4
snowd
Mean = 1.54
Var = 1.94
Frequency
Depth [m]
elev
Mean = 3.91e+03
Var = 9.15e+03
Frequency
Elev (m)
solar
Mean = 2.27e+05
Var = 1.82e+09
Frequency
Solar [WH/m2]
×10^5
aspect
Mean = 166
Var = 4.04e+03
Frequency
Aspect [deg]
×10^6
UTMN (m)
UTME (m)
×10^5
wind
Mean = 0.718
Var = 229
Frequency
Wind [Tbreak]

Senator Beck Basin -- SubCatchment:4

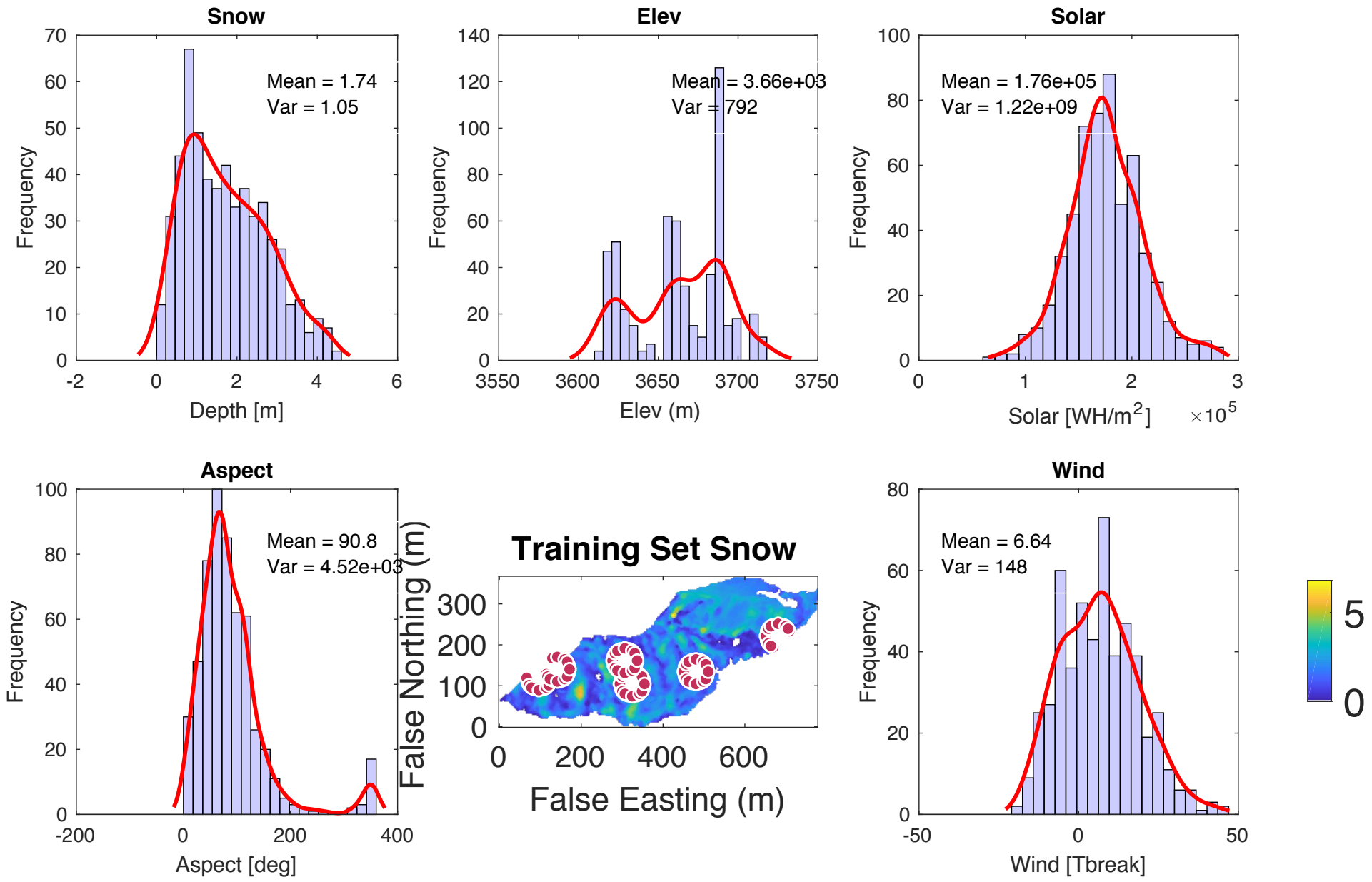


Senator Beck Basin -- SubCatchment:9




Senator Beck Basin -- SubCatchment:9
snowd
Mean = 1.92
Var = 0.916
Frequency
Depth [m]
elev
Mean = 3.59e+03
Var = 5.23e+03
Frequency
Elev (m)
solar
Mean = 1.32e+05
Var = 3.97e+09
Frequency
Solar [WH/m2]
×10^5
aspect
Mean = 96.5
Var = 8.21e+03
Frequency
Aspect [deg]
×10^6
UTMN (m)
UTME (m)
×10^5
wind
Mean = 4.51
Var = 194
Frequency
Wind [Tbreak]

Senator Beck Basin -- SubCatchment:9



Questions?