APPLICATION OF VARIABLE LAG VARIOGRAPHY ON DIRECTIONS DERIVED USING K-MEANS CLUSTERING OF SAMPLE PAIRS

ΕΓΕ 2019

15th International Congress of the Geological Society of Greece

Athens, 22-24 May 2019

Ioannis Kapageridis *PhD CEng CSci MIMMM*

Department of Mineral Resources Engineering

University of Western Macedonia

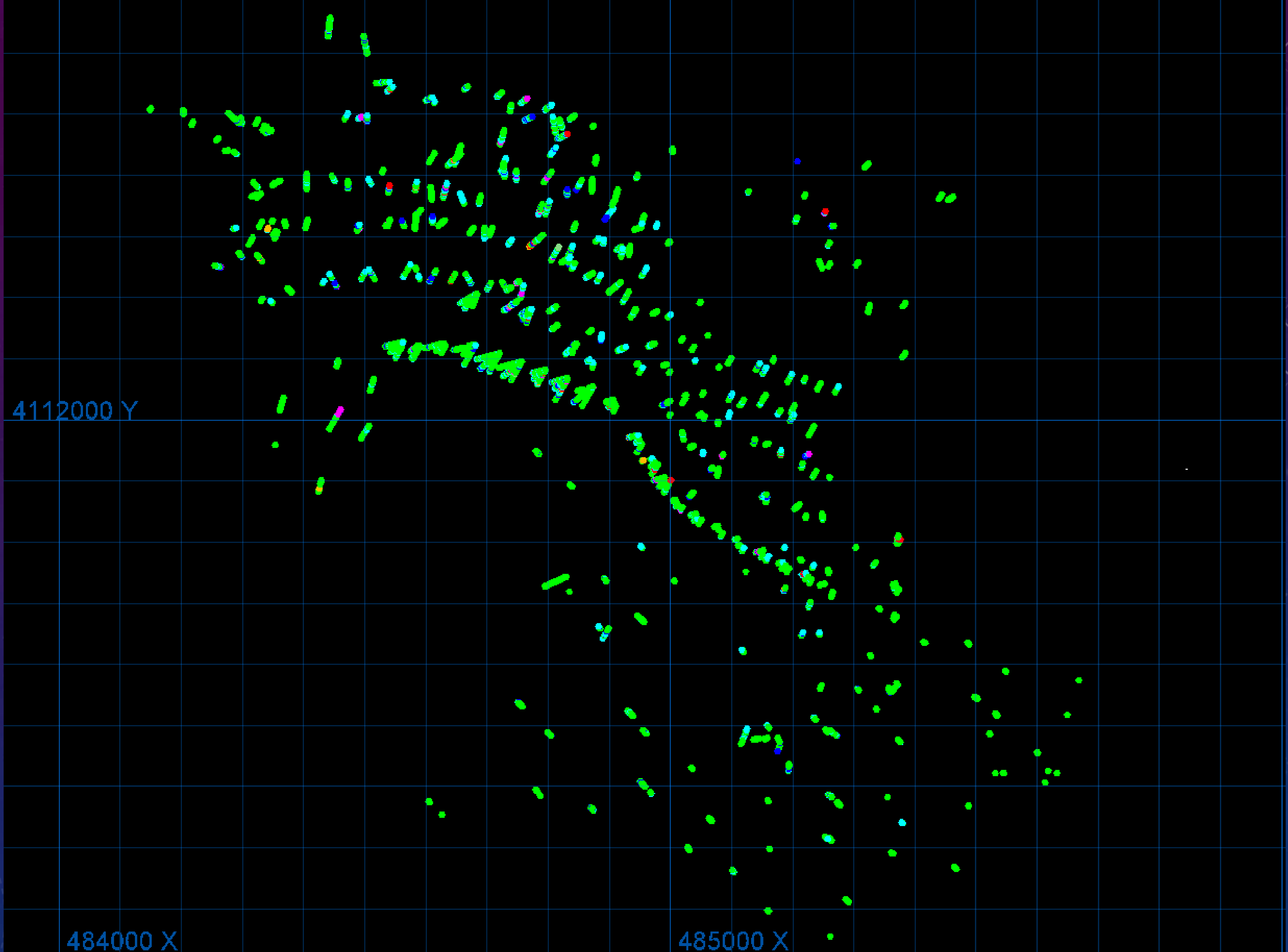
THE PURPOSE OF VARIOGRAPHY (STRUCTURAL ANALYSIS)

Variography aims to derive a coherent model for spatial variability of the considered mineralisation. This model can be used to:

1. Characterise the **spatial variability** of the mineralisation
2. Enable modelling of **change of support**
3. Help assess the effectiveness of **sampling spacing**
4. Obtain a weighting function for optimal local estimation (**kriging**)
5. Provide a structural function for the **simulation** of the mineralisation.

IRREGULARITY IN SAMPLING PATTERNS

- Drilling patterns, in mineral exploration programmes in particular, very rarely follow a strictly regular pattern.
- In some cases, this is due to practical issues causing a deviation from a constantly spaced pattern, while in other cases, it is the geometry of the targeted orebody envelope that requires a more flexible pattern to be followed.
- As exploration takes place in stages, in-fill drilling to increase the level of confidence in certain areas, also causes local changes in sample spacing.
- Drilling from underground workings is, in most cases, irregular and leads to extreme variation of sample spacing.

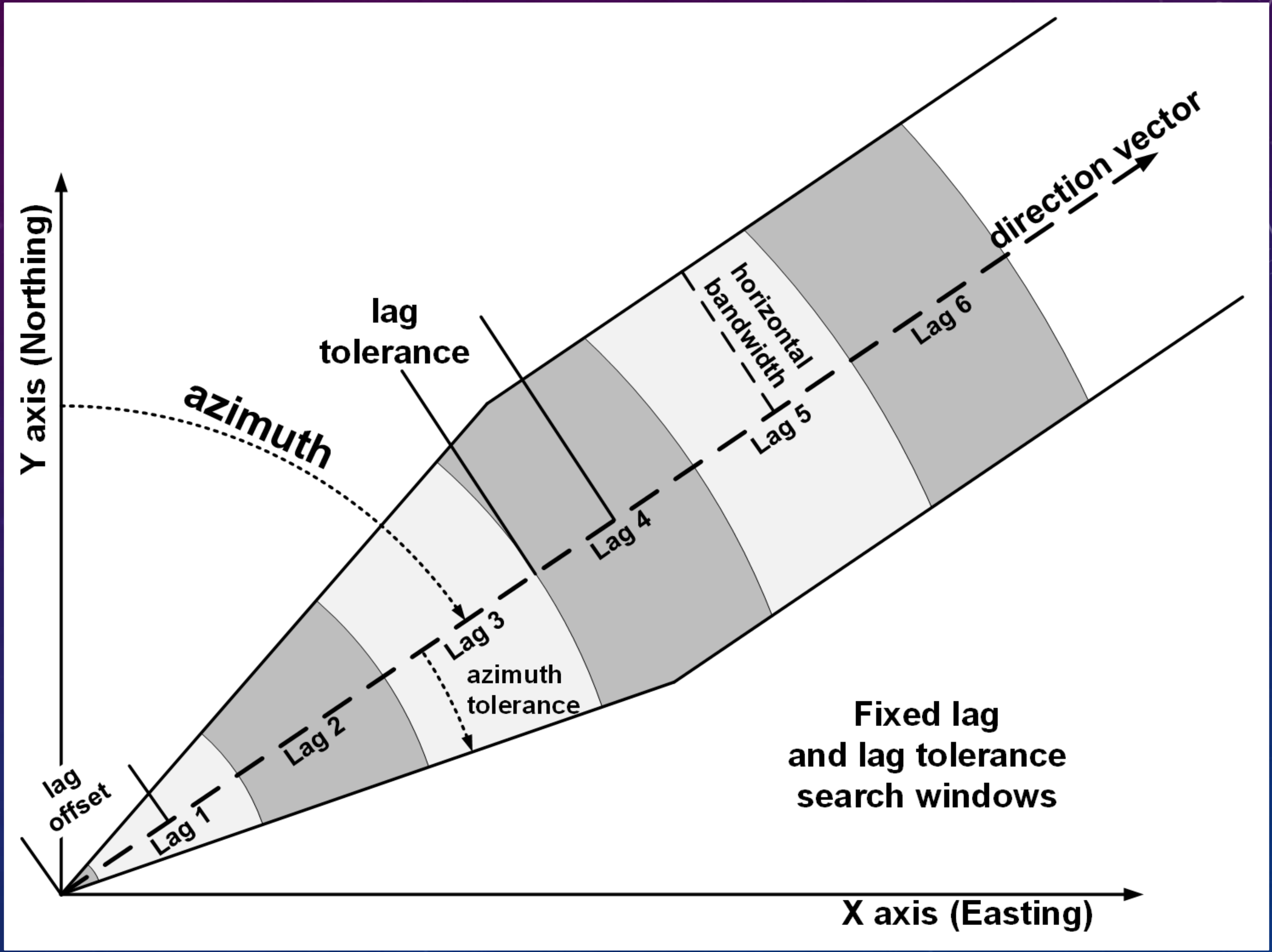

4112000 Y
484000 X
485000 X

THE PRACTICE OF VARIOGRAPHY

- Most geostatistical software packages, position a search area at each sample considered in order to locate a second sample to form a pair that will be used to calculate the value of the experimental variogram at various distances and directions.
- For a particular direction chosen, the **search area** is defined using some **direction tolerance** which can be controlled both **horizontally** and **vertically**.
- These tolerances are allowed to expand the search area as the separation distance increases up to a maximum distance (**bandwidth**) from the direction vector.
- This way, the search area begins as a cone of circular or elliptical section (depending on whether the azimuth and plunge tolerances are different), and then becomes a cylinder of similar section to the cone, once the maximum distance from the direction vector is reached.

THE PRACTICE OF VARIOGRAPHY

- The search area is divided into windows of different separation distances.
- The separation distances are commonly a multiple of a basic distance called the lag.
- A standard lag tolerance is applied throughout the search area and in all search windows, commonly defined as a percentage of the lag or explicitly.
- The sample at the head of the search area and the sample inside each search window (tail) form a pair that contributes to the particular experimental variogram point.


Y axis (Northing)
X axis (Easting)
azimuth
lag
tolerance
direction vector
horizontal
bandwidth
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
azimuth
tolerance
lag
offset
Fixed lag
and lag tolerance
search windows

FROM EXPERIMENTAL VARIOGRAPHY TO THE VARIOGRAM MODEL

- Once all possible experimental variogram points are calculated for different directions and separating distances, the experimental variogram points for a particular direction are displayed in a graph to enable fitting of an appropriate variogram model.
- Fitting of the variogram model is necessary to provide a variogram value for all possible distances, as required by the kriging system.
- The shape defined by the experimental variogram points can be quite hard to interpret and model, as the irregularity of the sampling pattern can lead to insufficient number of sample pairs at particular separation distances and the random inclusion/exclusion of extreme values.

THE CONCEPT OF VARIABLE LAG AND VARIABLE DIRECTIONS VARIOGRAPHY

- The problem of grouping pairs of samples in particular directions seems to be quite suitable for solving using a clustering algorithm like k-means.
- The main concept is that pairs are clustered to particular directions according to the vector they define and then they are grouped according to a criterion like the separation distance.
- This way, the practitioner does not have to spend time and effort in finding appropriate lags and tolerances that work in that particular direction.
- A separate clustering run will have to be performed for each direction considered.

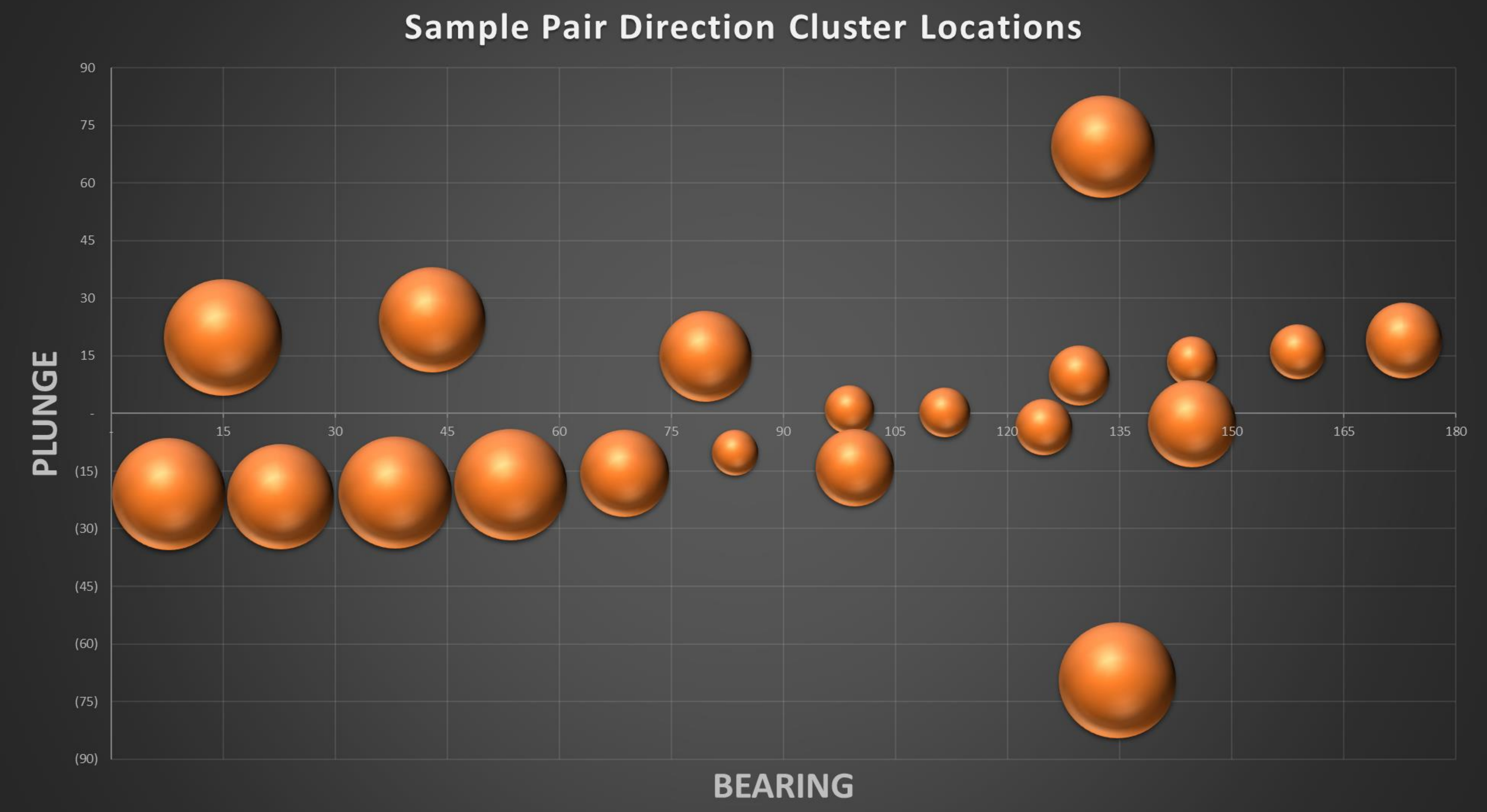
K-MEANS CLUSTERING

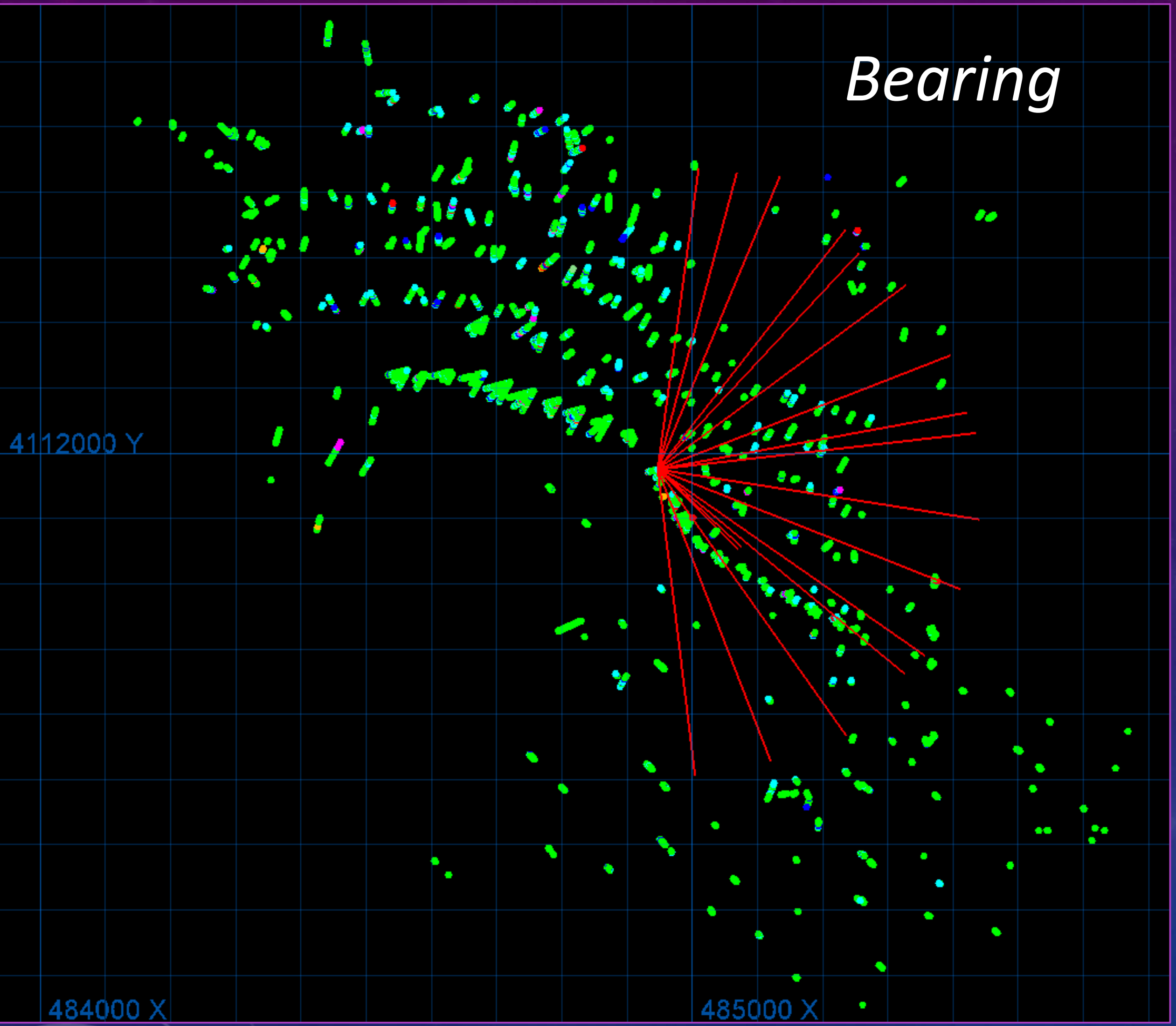
- k-means clustering is an iterative method originally used in signal processing, commonly used for cluster analysis in data mining.
- k-means clustering groups *n* observations into *k* clusters, with each observation assigned to the cluster with the nearest mean.
- Before any iteration, the clusters are initially centred on an equal number of observations using different methods.
- In the first step of an iteration, each observation is assigned to the cluster whose mean yields the least within-cluster sum of squares.
- The second step involves the calculation of the new means to be the centroids of the observations in the new clusters.
- The algorithm converges when there is no change in the assignments.

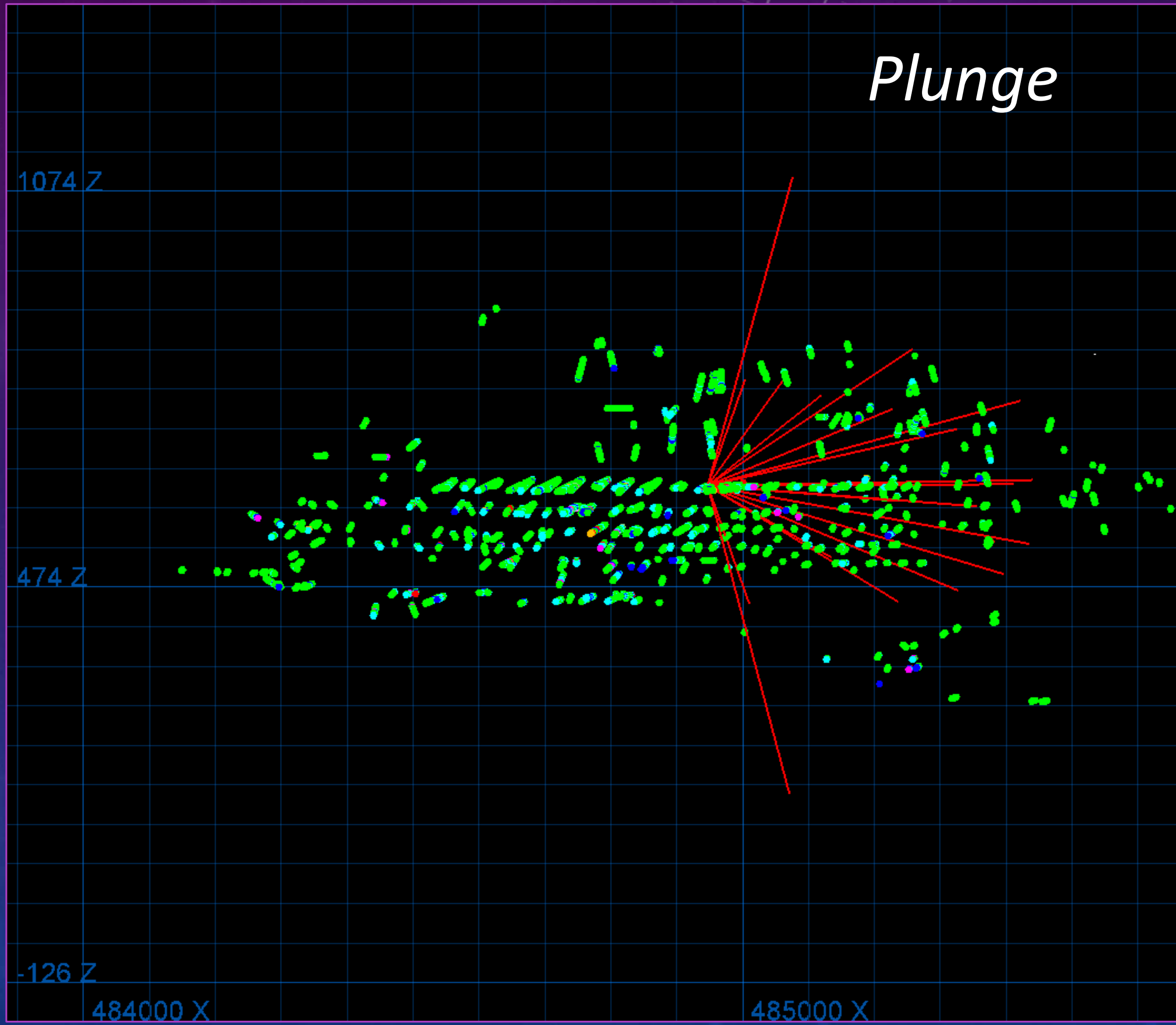
K-MEANS CLUSTERING OF VARIOGRAM DIRECTIONS

- In this study, samples from a single domain of a tungsten deposit were used.
- The drilling pattern was particularly irregular.
- Previous efforts on variography gave very inconsistent results – partly due to the sampling pattern irregularity and partly due to the high small scale variability of tungsten (nugget effect).
- Sample pairs were formed between all samples up to 300m of separating distance.
- The separating distance, bearing and plunge of the sample pair vector were recorded.
- Bearing and plunge of the sample pairs was used as criteria for clustering using k-means.
- 20 clusters were initially tried, with 5 of them giving a very low number of pairs and thus being excluded from further study.

Cluster Number	Bearing	Plunge	Maximum Distance	Pair Count
17	7.75	- 20.96	67.06	556,955
7	14.95	19.83	69.54	39,175
9	22.66	- 21.60	62.73	493,228
1	38.01	- 20.62	66.76	535,098
20	42.97	24.44	62.89	58,274
11	53.45	- 18.59	66.34	572,756
5	68.75	- 15.55	52.08	569,636
12	79.60	14.93	54.10	74,800
2	83.50	- 10.19	27.10	493,924
19	98.85	1.00	28.96	560,180
3	99.54	- 14.08	46.00	287,182
15	111.59	0.25	29.49	679,832
13	124.86	- 3.61	33.38	435,226
16	129.59	9.92	35.46	478,454
8	132.76	69.46	61.02	9,361
6	134.70	- 69.53	68.85	9,145
14	144.67	13.47	29.54	527,910
18	144.68	- 2.64	51.78	241,134
4	158.82	16.03	32.57	570,932
10	173.04	18.99	44.90	566,458

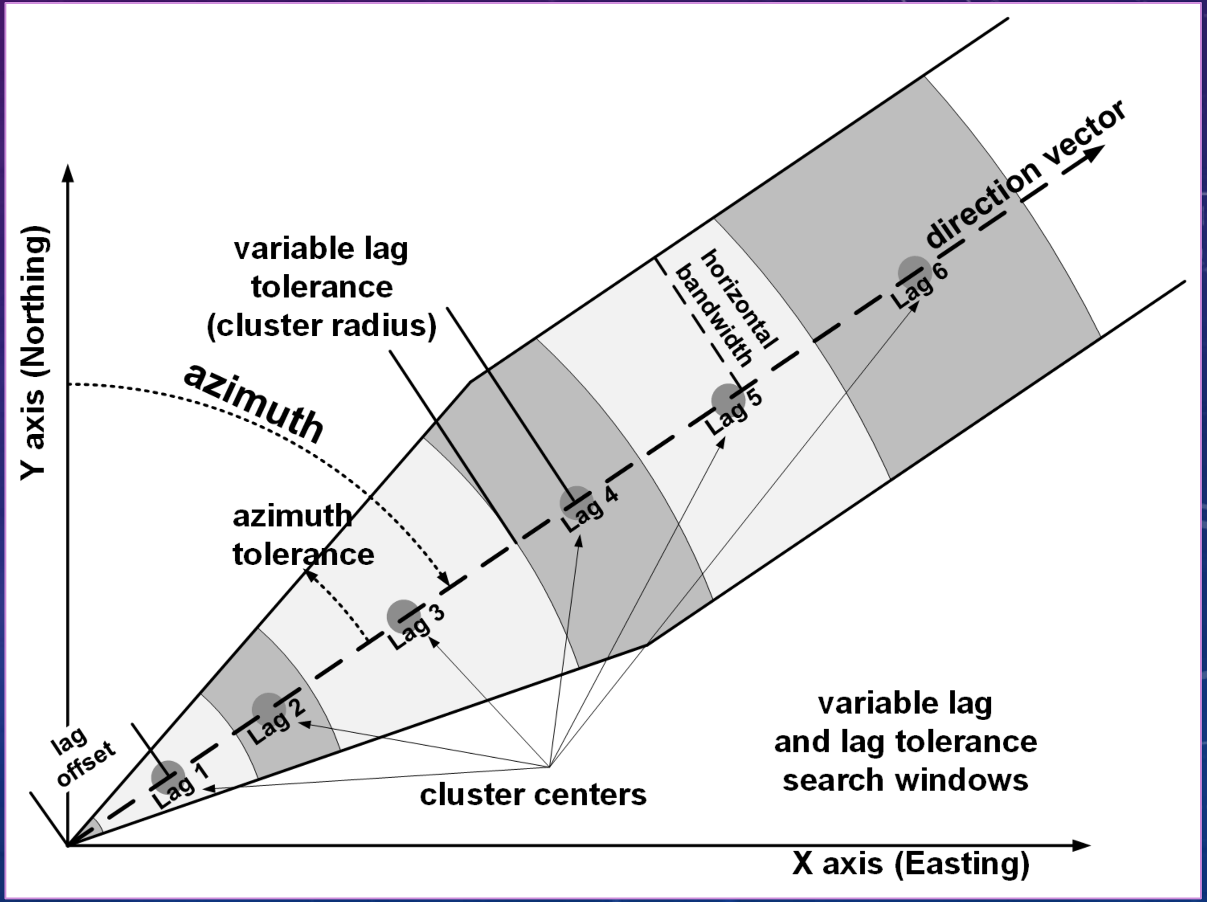

Sample Pair Direction Cluster Locations
PLUNGE
90
75
60
45
30
15
-
(15)
(30)
(45)
(60)
(75)
(90)
BEARING
15
30
45
60
75
90
105
120
135
150
165
180


Bearing
4112000 Y
484000 X
485000 X


Plunge
1074 Z
474 Z
-126 Z
484000 X
485000 X

VARIABLE LAG VARIOGRAPHY

- VLV has been introduced by the author as a way to automatically adjust lag parameters to match the spatial distribution of sample locations and improve the resulting experimental variogram.
- Sample pairs selected for a particular direction using standard criteria or the clustering process described in the previous step, are grouped into clusters based mainly on the 3D distance between their samples.



VARIABLE LAG VARIOGRAPHY

Variography parameters using k-means clustering in VLV are represented by the resulting clustering information:

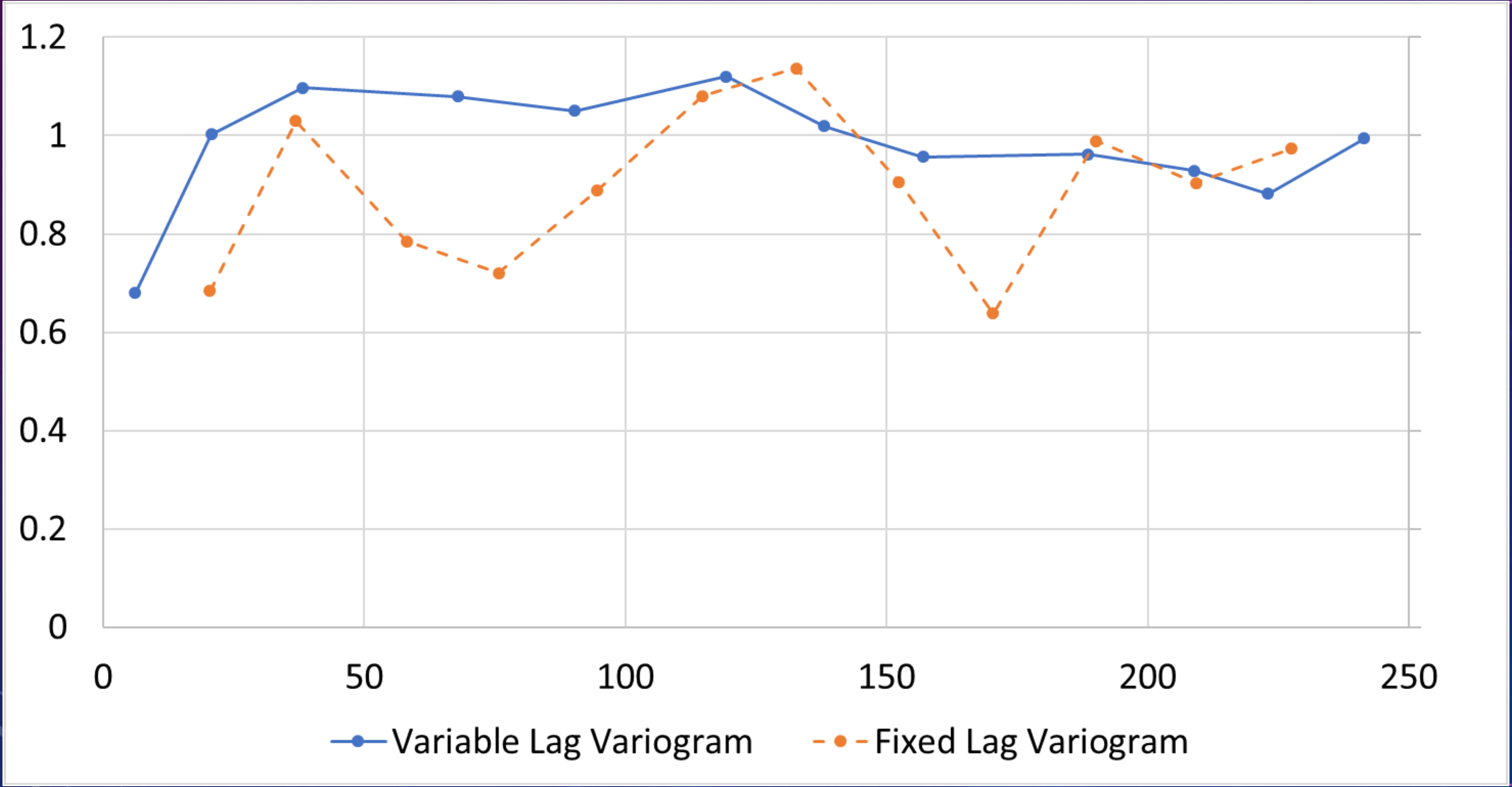
- ***Lag offset***: the average separation of the first cluster (first variogram point).
- ***Lag***: the average separation of each cluster (each variogram point) - different for each variogram point.
- ***Lag tolerance***: the maximum distance of the pairs classified in a specific cluster from that cluster's center - different for each variogram point, not a fixed value.
- ***Pair count***: the number of pairs classified in each cluster.

COMPARISON WITH STANDARD (FIXED LAG) VARIOGRAPHY

- The experimental variogram was calculated in the 15 directions found using k-means clustering.
- Both standard (fixed lag) and variable lag variography was applied in these directions.
- Standard variography benefitted from the use of directions found using clustering.
- In all cases, variable lag variography produced at least equally good results with standard variography, while in some cases the improvement in the shape of the variogram graph was significant.

FIXED LAG VS VARIABLE LAG EXPERIMENTAL VARIOGRAM

(*cluster #11, bearing 53.45º, plunge -18.59º*)



Point	VLV	Fixed Lag
1	7694	294
2	5244	723
3	3172	1707
4	7697	3356
5	9707	3661
6	9633	3772
7	5457	2826
8	6034	3618
9	4084	873
10	3858	2326
11	3874	2263
12	3180	2065

CONCLUSIONS

- The results from calculating experimental variograms using both approaches have shown that the application of variogram direction clustering and variable lag variography can relieve the practitioner from the trouble of finding well informed variogram directions and an appropriate lag setup.
- At the same time, the proposed solution based on k-means clustering, produces experimental variograms which are smoother and easier to interpret in cases of irregular sampling patterns.
- Current and further work includes development of an automatic cluster validation and selection method, a faster system of sample pair formation, and direct modelling of sample pair variogram clusters based on both directional and separating distance information.

Thank you for your attention!