

Tutorial – Clustering & Lithology Module



CONTENTS

Part 1 - Cluster analysis2
1.1 Cluster analysis of log data2
1.2 Cluster methods available in CycloLog2
1.3 The clustering menu and toolbar3
1.4 K-Means Clustering4
1.5 Native Clustering7
1.6 Hierarchical Clustering (Constrained)11
1.7 Hierarchical Clustering (Unconstrained)14
Part 2 - The lithology library
2.1 Working with the lithology library18
2.2 Property grid of Histogram23
2.3 Displaying lithology patterns in a log24
2.4 Saving the Histogram24

Copyright © 2016 ENRES International BV. All rights reserved

No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without written consent of ENRES International. Under the law, reproducing includes translating into another language or format. As between the Parties, ENRES International retains title to, and ownership of, all proprietary rights with respect to the software contained within its products. The software is protected by copyright laws. Therefore, you must treat the software like any other copyrighted material (e.g., a book or sound recording). Every effort has been made to ensure that the information in this user guide is accurate. ENRES International is not responsible for printing or clerical errors. Information in this document is subject to change without notice.

CycloLog is a registered trademark of ENRES Energy Resources International BV INPEFA is a registered trademark of ENRES Energy Resources International BV ENRES is a registered trademark of ENRES Energy Resources International BV

Version September 2016

Part 1 - Cluster analysis

1.1 Cluster analysis of log data

Cluster analysis in general is a method for exploring ways of arranging multivariate data into groups, or clusters. For the well geologist, it is a way of analyzing log data into electrofacies, each electrofacies being a way of grouping together data points with a similar range of properties. The variables in the case of log data are the various logs, any number of which can be selected for inclusion in the analysis.

The simplest form of clustering is the classification of a single log into a number of 'bins' of equal size, usually used for a colour display of the GR log in a way that suggests the range of lithologies represented. However, this method takes no account of the way the data points are distributed among the arbitrarily defined bins, and this section of the tutorial describes how to use the more sophisticated clustering tools available in CycloLog.

1.2 Cluster methods available in CycloLog

Cluster analysis has a very wide range of applications in many different disciplines, and many different clustering methods exist. Of these, CycloLog offers two *non-hierarchical* methods, and one *hierarchical* method.

The two non-hierarchical methods in CYcloLog are:

- Native clustering
- ➢ K-means clustering

Native clustering is quite intuitive, resembling the way that the eye tries to find clusters among points on a 2-D plane. It looks for groups of points that are each closely clustered around some centre point, at the same time trying to find group centres that are as far apart from each other as possible.

K-means clustering seeks the best way of dividing the data-points among some number (K) of clusters such that each data-point belongs to the cluster with the nearest mean value.

Hierarchical clustering in CycloLog works from the bottom-up, starting by assigning each data point to a separate cluster, then progressively adding clusters together based on their similarity. (Other hierarchical methods work from the top-down, successively subdividing the data into a larger and larger number of clusters.)

It is not easy to suggest criteria for choosing one method or another, and the user is invited to experiment with all of them. The following gives introductory instructions for each method.

1.3 The clustering menu and toolbar

Cluster analysis functions are grouped together in the **Clustering** menu (in the main menu bar, at the top of the CycloLog screen):

Clu	stering
	K-Means Clustering
	Native Clustering
	False Color Display
1	Cluster Matrix
Er.	2-D Crossplot
2	3-D Crossplot
Đe.	Hierarchical Clustering

A small number of clustering functions can also be accessed from the **Clustering** toolbar. Position the cursor over the icons on the toolbar to reveal their functions.



1.4 K-Means Clustering

To run a K-means clustering:

➢ Go to the main menu bar and select: Clustering → K-Means Clustering. The following dialog box opens:

Clustering				×
Cluster Logs	Cluster Interval	Cluster Parameters		
Source				
Well:	Well B			\sim
Domain:	Depth domain			\sim
Available GR DT RHOB NPHI ILD ILD ILD ILD ILD CALI CLS_GR D-INPEF D-INPEF	(d) (h) A_GR Total A_GR Blue			▲
			OK Annu	lleren

- > In the **Cluster Logs** tab, select the **Well** and **logs** you wish to run a K-means clustering.
- > In the **Cluster Interval** tab, define the **Top** and **Bottom** depth of the interval to be analysed.
- In the Cluster Parameters tab, select the number of clusters (default = 10): this should be larger than the expected final number of clusters.
- Click **OK** to run the analysis.

The analysis runs, and the following graph is displayed:



The graph shows an evaluation of the outcome for numbers of clusters up to the number you defined (12 on the example above).

- In the Number of clusters drop-down list, select the number corresponding to the point at which the graph begins to flatten (from left to right). In the illustrated example, you might choose 6 or 7.
- Click OK.
- > A cluster log called **K-Means** is added to the workspace and can now be displayed.
- ➤ To display the colour fill of the K-means cluster log, right-click over the K-means data pane log and select Display → Fill → Right → Log → K-Means. OR open the Properties toolbar (see next page, below). In the Properties pane, use the Log Fill options. Choose Log in the Fill with entry and select the cluster log.

GR - Depth do Depth domain Well B	DT - Depth do C	Pepth domain Well B	Epth domain Well B
3340 -	3340 - W	3340 -	3340
3360	3360 -	3360 -	3360
3380 -	3380	3380 -	3380
3400	3400	3400 -	3400
3420	3420	3420	3420
3440	3440	3440	3440
3460	3460	3460 -	3460
3480	3480	3480	3480
3500	3500	3500 -	3500
3520	3520	3520	3520
3540	3540	3540 -	3540

The **Cluster Matrix** can be displayed and modified, as described for the Native method below.

Pro	operties		>
			~
	2↓		
Ξ	Log appearance		/
	Line colour	000000	
	Line style	Solid	
	Line thickness	1 pixel	
Ξ	Log Fill		
	Fill side	Right	
	Fill with	Log	-
		K-Means	
	Color palette		
Ξ	Horizontal Scaling		
	Minimum:	0.000000	
	Maximum:	7.000000	
Ξ	Additional settings		
	Depth bar	True	
	Grid	False	
	Grid line color	c0c0c0	
	Grid line style	Dotted	

1.5 Native Clustering

To run a Native clustering:

➢ Go to the main menu bar and select: Clustering → Native Clustering. The following dialog box opens:

Clustering		×
Cluster Logs	Cluster Interval Cluster Parameters	
Source		
Well:	Well B	\sim
Domain:	Depth domain	\sim
Available GR DT RHOB NPHI ILD ILD ILD ILD ILD CALI CLS_GR CLS_GR D-INPEF D-INPEF	logs (d) (h) =A_GR Total =A_GR Blue OK	Annuleren

- In the Cluster Logs tab, select the well and the logs to be included in the cluster analysis (see example in the figure above).
- > In the **Cluster Interval** tab, define the **Top** and **Bottom** depth of the interval to be analysed.
- > Open the **Cluster Parameters** tab:

Clustering		Х
Cluster Logs Cluster Inte	rval Cluster Parameters	
Clusters		
Number of clusters:	7	
Decimation:	5	
Parameters		
Mass:	0.2	
Separation:	0.9	
Merging threshold:	0.02	
Options		
Move to center of	gravity	
	OK Annuler	en

- > For an initial run, accept the default parameters and click **OK**.
- > The analysis runs, resulting in a new **Cluster Log**, which is added to the workspace.
- It is a good idea to rename the cluster log: right-click over its name in the workspace, select
 Rename, and type the new name, then press Enter.
- > To view the cluster log, double-click on its name in the workspace.
- In the figure below, the cluster log has been coloured. Open the Properties pane and use the Log Fill options. Choose Log in the Fill with entry and select the cluster log.

GR - Depth do Depth domain Well B	DT - Depth do Depth domain Well B	Pepth domain Well B	Depth domain Well B
3340	3340 - W	3340	3340
3360	3360	3360 -	3360
3380 -	3380	3380 -	3380
3400	3400	3400	3400
3420	3420	3420	3420
3440	3440 -	3440	3440
3460	3460 - W	3460	3460
3480	3480	3480	3480
3500	3500	3500	3500
3520	3620 -	3520	3520
3540	3540	3540	3540

To view the numerical properties of the clusters:

- > Open the **Cluster Matrix** either from the Clustering menu, or from the Clustering toolbar.
- The cluster matrix shows the mean log values that define the centreof each cluster; this may be helpful for interpreting the geological or petrophysical significance of each cluster.

Sourc	e						014
Well:	: Wel	В			```	/ L	OK
Dom	ain: Dep	th domain			```	/	
Log:	Clus	ter log			```	/	
Matrix	(
	Cluster	GR	DT	RHOB	NPHI	Color	Descripti
\checkmark	0	30.55	85.71	2.42	24.44		
\checkmark	1	33.18	77.30	2.49	16.02		
\checkmark	2	40.98	75.72	2.50	19.73		
\checkmark	3	51.95	71.46	2.67	14.73		
\checkmark	4	57.39	77.82	2.42	17.03		
\checkmark	5	88.65	68.52	2.59	13.30		
\checkmark	6	94.37	72.55	2.49	20.19		
<							>

The **Cluster Matrix** can be used to analyse and modify the cluster results:

- > To change the colours assigned to a cluster, click on a colour in the **Color** column and select the desired colour.
- To add a text description to a cluster, click in the **Description** column of the cluster matrix, and type in your text.
- > The **Export** button can be used to export the cluster matrix to an ASCII file (*.cls).
- ➤ Note that the cluster log results can be used to colour fill other logs listed in the workspace (except special logs). Open a log from the workspace and make sure that it is the active log. Open the Properties pane and select Fill with → Log and subsequently select the cluster log from the drop-down menu.

1.6 Hierarchical Clustering (Constrained)

There are two different approaches to **Hierarchical Clustering** in CycloLog: **Unconstrained**, and **Constrained**. (*Examples are illustrated below*.)

- In the Constrained method, the analysis is forced to maintain the original order of the samples, so that the resulting dendrogram reflects the similarity between adjacent samples. This method may be also be called stratigraphic clustering.
- In the Unconstrained method, the analysis is not restricted to maintain the original order of the samples. Samples from the top of the well could from part of the same cluster as samples from the bottom of the well.

To run Hierarchical Clustering:

➢ Go to the main menu bar and select: Clustering → Hierarchical Clustering. The following dialog box opens:

Hierarchical (Clustering			×
Cluster Logs	Cluster Interval	Cluster Param	eters	
Source				
Well:	Well A			\sim
Domain:	Depth domain			\sim
Available I GR DT RHOB DRHO ILD CALI CLS_GR D-INPEF, D-INPEF, D-INPEF, D-INPEF,	(d) (h) A_GR Total A_GR Blue A_GR Brown A_GR Green A_GR Purple A_GR Purple-2			~
			OK	Annuleren

- In the Cluster Logs tab, select the Well and Logs.
- > In the **Cluster Interval** tab, define the **Top** and **Bottom** depth of the interval to be analysed.
- > Open the **Cluster Parameters** tab:

Hierarchical C	lustering			×
Cluster Logs	Cluster Interval	Cluster Parameters		
Method				
Cons	trained			
	nstrained	Preprocessing st	ep: 250	
Distance Fi	unction			
Euclid	lean			
() Manh	attan			
Linkage				
○ Neare	est-neighbour			
○ Furth	est-neighbour			
O Avera	age			
Centr	oid			
Ward				
			OK	Annuleren
				r an an an an an an an

- Select **Constrained**.
- For an initial run at least, accept the default Distance Function and the default Linkage method.
- Click **OK** to run the analysis.
- > Open the resulting **Hierarchical Clustered Log** that appears in the workspace.

The following figure shows a constrained hierarchical clustering of the same stratigraphic interval that was used for the Native and K-means examples earlier in this section (i.e., Well B and logs GR, DT, RHOB and NPHI). The display is in the form of a dendrogram, in which the horizontal axis shows the degree of similarity of adjacent data-points, from most similar (left) to least similar (right).

💮 CycloLog 2016 - Base Tu	torial_Correlation panel.clg	- 0	×
<u>File Edit View Output</u>	ut <u>L</u> ayout <u>C</u> lustering <u>A</u> n	rsis <u>C</u> alculations <u>P</u> etrophysics Batch Operations <u>T</u> ools <u>W</u> indow <u>H</u> elp	
i 🗅 💕 🖬 🗠 🗠 🎒	🛕 🛯 🛤 💷 🕷 🏕	u 🐵 🚛 💎 📢 🖓 📅 🔄 v = 🚃 🚛 📑 🖬 🖬 🖬 🗰 🖄 👘	<u>N</u> a →
NBS -	3297.51 Top		- =
Properties			
<		🗮 🖶 GR - Depth domain - Well B	×
		Depth domain Depth domain	
S SII Z↓		Wein D Wein D	
E Log appearance	I a an Albania		
Display mode Dendrogram	True		
E Log Bill	100		
Fil	False		
Horizontal Scaling			
Minimum:	0.200000		
Maximum:	1000.000000		- 10
Additional settings			
Depth bar	True		
Grid	False		
Grid line color			
Cluster column	Dotted		
Cluster column	False		
Cidator Coldmin	1 0.00		
		3440 - 34400 - 34400 - 34400 - 34400 - 3440 - 3440 - 3440 - 3440 - 3440 - 3440 -	
		3460	
Maximum:			
The maximum value for the	horizontal scaling.		~
			_
🚺 Information 🥖 Logo	ging Sheet 📋 Memo Sheet		
Depth: 3459.978 meter / 113	51.634 feet Hierarchical Clus	ed Log: 0 Normal Normal CAP NUN	M SCRL

The display of this Hierarchical Clustered Log has been made easier to read by:

- > Make sure the Hierarchical Clustered Log is open and the active pane.
- > Open the **Properties** pane.
- > In **Log appearance**, change the **Display mode** to *Logarithmic*.
- In Horizontal Scaling, change the minimum to 0.2 (you may test what is the preferred view, by using a lower or higher value). Don't change the Maximum value.

It is now easier to see the relationships between samples. Note, for example, that the pair of adjacent samples that are most different from each other are below 3460m (because they belong to the two largest clusters). In this way you can determine a hierarchy of relationships between adjacent samples.

For additional options, see section 18.6 Hierarchical Clustering in the CycloLog Help manual.

Version September 2016

1.7 Hierarchical Clustering (Unconstrained)

In Unconstrained Hierarchical Clustering, the clustering is unconstrained by the relative stratigraphic (depth) positions of the data points.

CycloLog does the clustering in two stages (to avoid unnecessarily long computing time). The first 'preprocessing' stage is a K-means clustering, which provides the part-processed data on which the hierarchical clustering subsequently operates. The number of pre-processing steps can be defined by the user.

- > Run as for the Constrained method, except:
- Select **Unconstrained** in the **Cluster Parameters** tab of the Hierarchical Clustering dialog box.
- > Either accept the default (=250) **Pre-processing steps**, or change to your preferred number.
- Click OK.
- CycloLog performs the clustering and adds a new Hierarchical Clustered Log to the workspace.
 (You may change the name by right-click with the cursor on the log's name in the workspace and then select Rename).
- > Open the new log by double-clicking on its name in the workspace.

Note that the target number of clusters is *not* defined before the analysis. Initially, the display will show the data grouped into only two clusters, like the right log data pane in the illustrated example:



To change the number of clusters, right-click with the cursor over the relevant display pane, and click on **Set Clusters** to open the following, initial, Clusters dialog window:

Clusters	×
Dendrogram Cluster Matrix	
1	2
Display	Clusters
Normal	Number of clusters: 2
	Reduster
	OK Annuleren

- > This window shows the **dendrogram** for the clustered data, with only two clusters (initially).
- Reset the Number of clusters to 10.
- > Click on **Recluster**. The same data organized into ten clusters is shown below.



> The dendrogram shows the relative similarity/dissimilarity between the clusters. In the figure below, the curves of the 2 and 10 number of clusters are shown in the data panes on the right.



To find out which cluster each data-point has been assigned to, place the cursor over the **Hierarchical Clustered Log** pane, hold down the left mouse button, and move the cursor up or down to the required depth. Press the key-board letter L and the cluster number will be shown (see blue line in image above). It also will be shown in the **Status Bar** at the bottom of the CycloLog screen.

You can also access the **Cluster Matrix** from the Cluster Matrix tab in the Clusters dialog window. This shows the mean log values for the currently defined set of clusters, which may help in interpreting the geological or petrophysical significance of the clusters.

In the **Cluster Matrix**, you also can assign different colours to each cluster. Select the cluster you wish to add colour to and click on the **Color** button. Alternatively, colouring of clusters can also be added from the Dendrogram tab. The square boxes at the ends of the dendrogram can be double clicked and a colour can be added. Note that a coloured Hierarchical Clustered Log can also be used to colour other logs.

Part 2 - The lithology library

2.1 Working with the lithology library

The lithology library can be accessed through the Histogram pane of a log. To open the Histogram:

- \succ Right-click with the mouse button on the log pane and select Analysis \rightarrow Histogram.
- Alternatively, you can access the histogram from the right-click menu on a log in the workspace (select View Histogram).
- > The Histogram for the selected log is shown in a pane.



By default, vertical lines representing the minimum and maximum value of the data and the cut-off for toning are displayed. To change their values, they can be dragged along the log range. The interval defined by the minimum and maximum is used to calculate the relative frequency. Relative frequency is recalculated based on a newly defined minimum and maximum when **Apply Scaling** is set at **True** in the **Properties** pane. The Cut-off value is used for toning of the log which the is displayed in the histogram. Toning can be applied via the **Properties** pane (see next image).

Version September 2016

Prop	Properties ×					
	© I A					
Ξ	Histogram Appearance					
	Title	Histogram - GR - Depth domai				
)	X - Label	GR (-)				
	Y - Label	relative frequency (%)				
1	Show Grid	True				
0	Cumulative	False				
	Histogram Parameters					
	Top Depth	3200.000000				
	Bottom Depth	3663.800049				
1	Number of bins	100				
\Box	Scaling					
:	Show Boundaries	True				
[Horizontal scaling					
	Minimum	7.136564				
	Maximum	130.440521				
	Cut-off	20.616739				
	Apply Scaling	False				
	Reset Scaling	False				

The log range in the histogram can be divided into subranges. Each subrange can then be assigned to a lithology. To make subranges, add pegs on the horizontal axis of the histogram:

- To set a peg, move the mouse pointer below the horizontal axis of the histogram and double click the left mouse button. A green coloured peg will be added.
- Press the left mouse button on a peg if its position needs readjusting. This allows dragging it along the log axis.
- > Right click on a peg to remove the peg and delete the left or right range.
- When you define or delete a range by setting or removing a peg, the legend of the histogram is updated accordingly.

To each subrange, a lithology pattern can be assigned:

Click on the lithology pattern in the legend of the histogram. This will open the Lithology Library:

Pattern	Back	Fore	Name	Code	Group	Description	^	
			Carbonate	Rock01	Carbonates			
			dolomite	Rock02	Carbonates			Add Edit
			m-c Sandstone	SST01	Clastics	Medium-Coarse Sandstone		Delete
			m-f Sandstone	SST02	Clastics	Medium-Fine Sandstone		Batch Ad
			Mud	Rock29	Clastics	Mud-shale		
			Rock03	Rock03	Other			
			Rock04	Rock04	Other			
			Rock05	Rock05	Other			
							~	

- Select a lithology pattern in the lithology library.
- > Click **OK** to apply the pattern to the selected subrange in the histogram.
- See example below after having defined a lithology for each subrange.



Version September 2016

Within the lithology library, lithology patterns can be added, edited and deleted:

- > Open the **Lithology Library** by double clicking on a pattern in the legend of the histogram.
- > To create a new lithology pattern, click Add.
- > The **Create Lithology Unit** dialog box opens:

Create Litholo	gy Unit	×
Name	1	Code
Description		
Group	Other ~	Targets: Log analysis Core analysis Project specific
Pattern	Load	Background V
		OK Cancel

- Insert a Name and Code for the new pattern.
- > Adding a **Description** is optional.
- You can assign the new pattern to a lithology Group (The groups available in CycloLog are: Carbonates, Chemical, Clastics, and Other).
- Selecting a **Target** for the new pattern is optional.
- ► Load a pattern EITHER from the default lithology library (navigate to ProgramData → Enres International → CycloLog → Lithology → Library) OR a pattern of your own. In the latter case, the pattern file can have either the "bitmap" file format (*.bmp) or the "tagged image file" format (*.tif, *.tiff).
- > Adjust **Back** and **Foreground** colour.

Create Litholo	gy Unit	×
Name	MF-SST	Code 1001
Description	Medium to fine grained sand	
Group	Other 🗸	Targets: Log analysis Core analysis Project specific
Pattern	Load	Background V
		OK Cancel

- Click **OK**; the new pattern is saved to the Library.
- To edit an existing litholology pattern in the Lithology Library, click on the Edit button. Use the same workflow as outlined above for creating a new lithology pattern.
- To delete a lithology pattern from the Lithology Library, select the pattern and then click on the Delete button.

2.2 Property grid of Histogram

- The property grid of the Histogram can be accessed by opening the **Properties** bar (right click over the toolbar area and select Properties, if it is not already open).
- In the Histogram Appearance node, the Title of the histogram, and the X- and Y- Label of the histogram axes can be edited. Show Grid and the Cumulative display of the histogram can be turned on or off by selecting *True* or *False*, respectively.
- In the Histogram Parameters node, the analysis interval of the Histogram can be adjusted (thereby not changing the displayed interval of the log itself). Therefore, edit the **Top Depth** and **Bottom Depth** values. The **Number of bins** (default = 100) can be specified.
- In the Scaling node, you can choose to Show Boundaries (i.e., the Minimum, Cut-off and Maximum vertical lines in the Histogram pane). By default, Show Boundaries is set at True.
- In the Horizontal scaling node, interval defined by the Minimum and Maximum is used to calculate relative frequency for the data points within the defined values. Set Apply Scaling to *True* to recalculate relative frequency based on the newly defined data interval. You can reset this scaling by setting Reset Scaling to *False*. The Cut-off value is used for the log toning.

Pro	Properties X				
		~			
•	Å ↓				
Ξ	Histogram Appearance				
	Title	Histogram - GR - Depth domai			
	X - Label	GR (-)			
	Y - Label	relative frequency (%)			
	Show Grid	True			
	Cumulative	False			
Ξ	Histogram Parameters				
	Top Depth	3200.000000			
	Bottom Depth	3663.800049			
	Number of bins	100			
Ξ	Scaling				
	Show Boundaries	True			
	Horizontal scaling				
	Minimum	7.136564			
	Maximum	130.440521			
	Cut-off	20.616739			
	Apply Scaling	False			
	Reset Scaling	False			
		-			

2.3 Displaying lithology patterns in a log

To display the lithologies defined for a log (based on its histogram):

- Open the log from which the histogram and lithologies were generated by double clicking on it in the workspace tree.
- Go to, or open, the **Properties** pane of the log (in the example below, the GR log has been selected).
- > Go to the Log Fill node and select Fill with → Histogram.
- > The Log is filled with the lithology patterns defined above, in part 2.1.



The display of the lithology in the log pane is interactive: if the pegs are moved in the histogram, the log fill display is updated directly.

2.4 Saving the Histogram

- > The Histogram will be saved when saving and closing the CycloLog project.
- A Histogram will not be added as a separate pane to the workspace. It can only be accessed via the active pane of a selected log.
- ➤ A Histogram cannot be exported.