Introduction and software manual for SG2PS

(Structural Geology to Post Script)

Platform independent

structural geological data processing software

Coded by Ágoston Sasvári

Graphical User Interface coded by Ali Baharev

MAY 2017

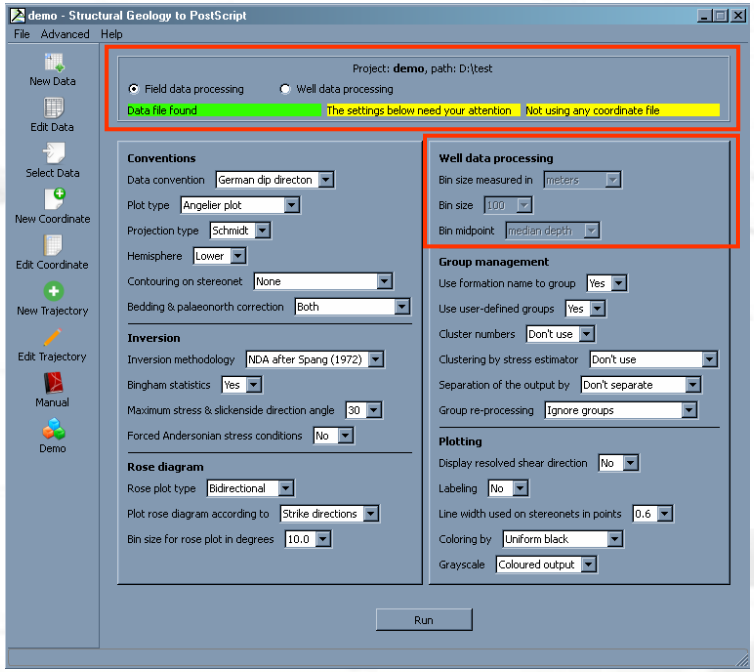
CONTENT I.

CONTENT II.

QUICK OVERVIEW

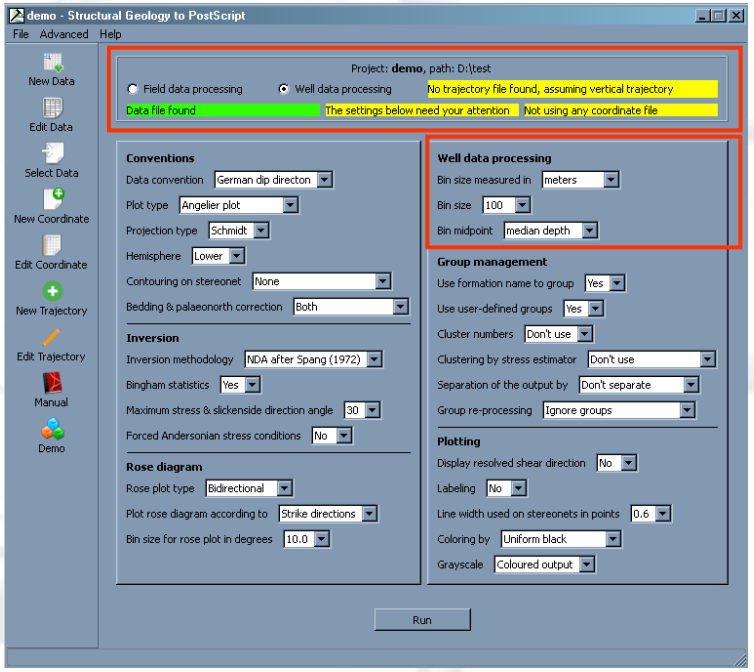
FIELD PROCESSING MODE

- **Formation boundary**, **lineation**, **plane**, **striae** and **SC** schistosity data can be processed
- If a **coordinate file** is present for your field data locations (as data point – coordinate – formation name sets), it will be taken into agreement
- Using average dips, a **tilting test** is carried out
- **Stereonets** and **paleostress** calculation results are exported in ***.eps** of ***.pdf** format



WELL PROCESSING MODE

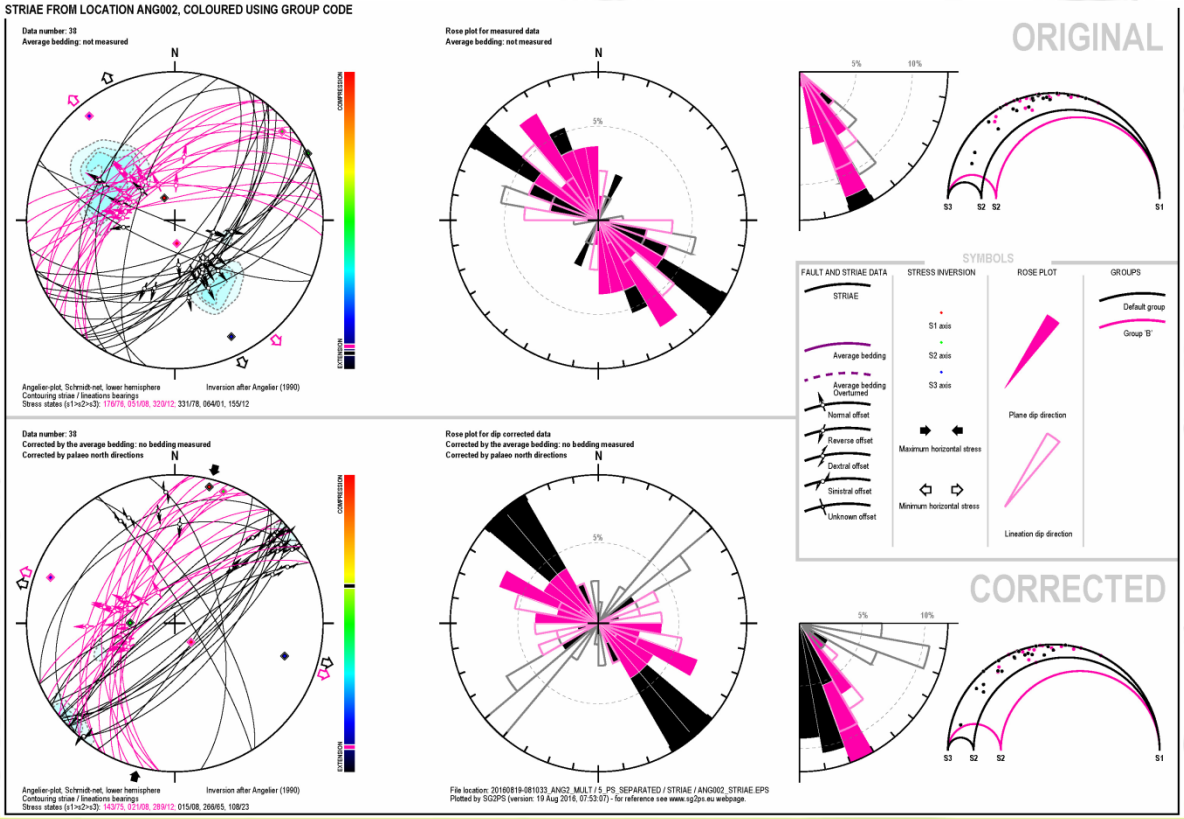
- XRMI / FMI processing results, such as **bedding**, **fracture** and **borehole break out** data can be processed; striae and SC data are ignored
- **Trajectory correction** (using a depth – coordinate or depth – azimuth – plunge records) on the can be applied in the same time as the tilting test; no coordinate file is processed
- **Stereonets**, and **faults probability graphs** are exported in ***.eps** of ***.pdf** format



QUICK OVERVIEW

FIELD DATA PROCESSING MODULE

- **SG2PS** is a **free** graphical application for **field** and **well** structural data evaluation and processing
- Easy to use **Graphical User Interface**, data processing takes just a few clicks
- Separate processing and plot for different **localities**, data **types** and data **groups**
- Displays structural data on **stereonets** and **rose diagrams**
- **Bedding correction** for tilting test available
- **K-means clustering** to identify data groups in an inhomogeneous data set
- Seven different **inversion** methodologies and fracture geometry statistics
- **Batch** processing for different input data sets
- Data **input in Excel** or any other spreadsheet application of your choice
- Output in Adobe Portable Document Format (**PDF**)



CAPABILITIES

- **Angelier** or **Hoeppner** plot, **Schmidt** or **Wulff** net, as well as **upper** or **lower** hemisphere projections are possible
- **Symmetric / asymmetric** rose diagrams with different **bin size** are available to plot dip **direction or strike of data**
- Contouring data set using four different possible inputs
- **Slickenside data correction** on two ways
- **Full automatic** data processing option
- Simplified data input and **auto completion** of input database
- Processing **by formations**
- Average data (e.g. average bedding) calculation using **regression**
- **Tectonic transport direction** display and **mean transport direction** calculation
- Displaying **multiple data groups** on the same stereonet
- Built in **k-means clustering** module
- **Stress tensor inversion** using
  - Turner (1953),
  - Spang (1972),
  - Michael (1984),
  - Angelier (1990),
  - Fry (1999),
  - Shan et al. (2003), and
  - Mostafa (2005) methodologies
- **Bingham directional statistics** for fracture sets
- **RUP** and **ANG** stress estimators display and slickenside clustering using them
- **Ideal movement** direction display
- Stress state display using **Mohr circle**

1) AVERAGE CALCULATION FOR STRUCTURAL TRENDS

- **Average** feature **dip / dip direction** calculation along well bore using **fix data number bin** or **fix interval bins**
- **Standard deviation** and **relative error** calculation for the processed bins
- **Feature frequency** calculation

2) DERIVATE CURVE TO SHOW OUT-OF-TREND ZONES

- Investigation of **out-of-trend** zones using **first derivate curve** of the average and frequency curves

3) PEAK DETECTION FOR DEFORMATION ZONES

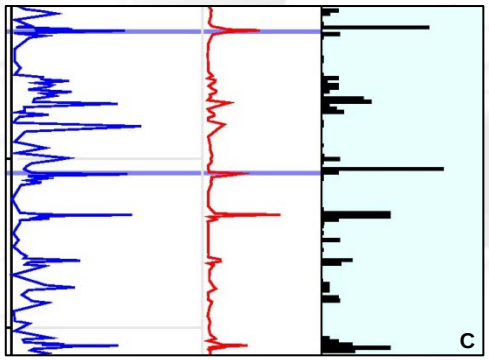
- **Changes** on average, frequency and derivate curve can indicate a zone of **deformation**
- **Peak recognition** of these curves can help to show the **reliability** of the assumed deformation zones

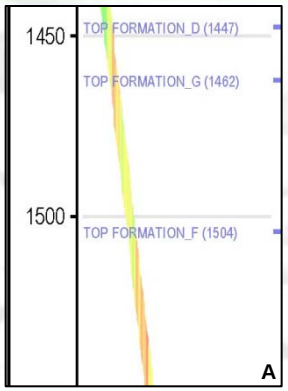
4) FAULTS PROBABILITY

- Abrupt **change** in feature geometry and/or frequency can indicate a **zone of deformation**;
- Using **changes** both of average and derivate curve, and **reliability** of those changes, **probability of fault** zones can be predicted

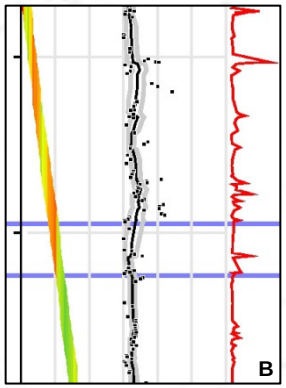
Directional data, average data, derivative curves, frequency curves and fault probability is **exported in *.csv** format

Measurements frequency (blue line) is plotted. Red line shows change in feature frequency as the first derivate of the frequency curve, indicative for faults. Reliability of peaks on the derivate curve are marked by colored dots. Faults probability (using average, frequency and derivate curves) is calculated for the well bore and displayed as black bars.

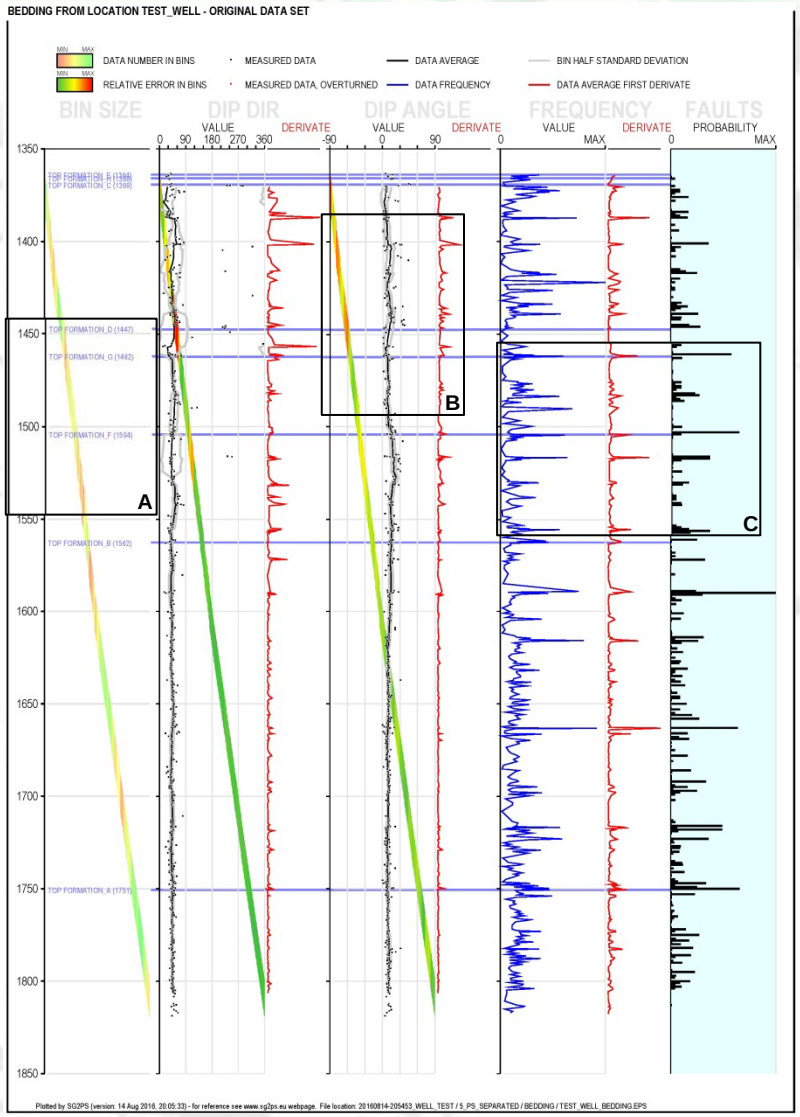




Bars showing relative data number in one bin: the higher the data number, the reliable the processing.



Measurements (dots), their moving average (black line) and standard derivate (gray line) plotted. Colored bars for relative error, and red line shows intensity of change in feature geometry as the first derivate of the average curve. Changes on the average and derivate curve are present as peaks, and their can indicate faults.



Final output, showing 1) bins size and their relative error, 2) dip direction trend and its change, 3) dip trend and its change, 4) feature frequency and its change, and 5) faults probability.

THE SOFTWARE

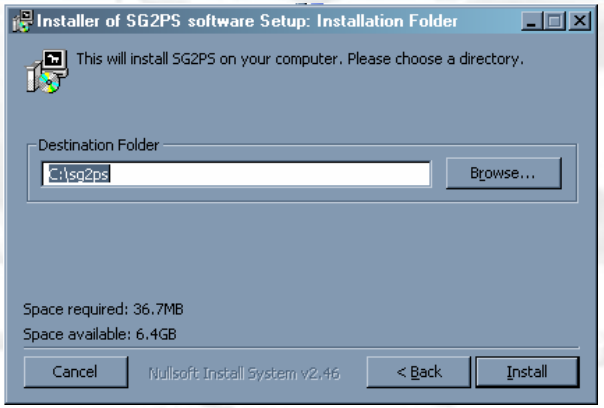
- Download SG2PS installer from the website **http://www.sg2ps.eu/download.htm** and install on unzip
- Start the Graphical User Interface (**SG2PS_GUI.EXE**)
- **Create** your data spreadsheet what you want to evaluate using the **graphical version** of the program, or **copy-paste into a text file** in the working folder – this will be the **input (*.rgf)** file
- You can store your **settings (*.set)** and **locality – coordinate – formation name pairs (*.xy)** in separated files
- For processing well data, you can use well trajectory (*.trj) file for borehole deviation correction either in depth – azimuth – plunge or in depth – X – Y data format
- A **project folder** will be generated for the text and post script or pdf outputs
- Evaluation results will be displayed on the **screen** in text format
- If everything goes well, **couple of thousand data** from **some hundred data points** could be evaluated and plotted in **less than one minute**
- To remove **evaluation results**, delete the project folder
- To remove **SG2PS**, just delete the program folder

THE MANUAL

- If you are interested in the software capabilities, or want to have a quick overview, read **INTRODUCTION** chapter; if you are interested is the result plot details, take a look at the **WELL AND FIELD DATA GRAPHICAL OUTPUT** chapter as well
- If you want to use the software to evaluate your own data, please have a look at the **INPUT FILE FORMAT** chapter to became familiar with the data input requirements
- Chapter **PROCESSING AND DISPLAY** gives a slightly detailed overview about the software capabilities presenting well data, and displaying different plot, net and hemisphere types as well as short description of used inversion methodologies
- **BACKGROUND** chapter describes all of the details what can help you to understand the source code. Please find here a brief summary about the different net calculation, Bingham statistics, clustering and inversion methodologies. This chapter don't repeat the published solutions, it is just a brief summary how to code them. Background theory and mathematics are published in the referred papers
- Feel free to use this software at any time, but please have a look at the **COPYRIGHT**

INSTALLATION

- Download the install file from the **www.sg2ps.eu/download.htm** webpage and run '**sg2ps_setup.exe**'
- Choose the directory where you want to install the application and press '**Install**'
- If you are not allowed to run self extracting archives, simply download the installer in ***.zip** format and extract in the destination folder



RUN – FILE MENU

- Start '**sg2ps_gui.exe**' to run the software

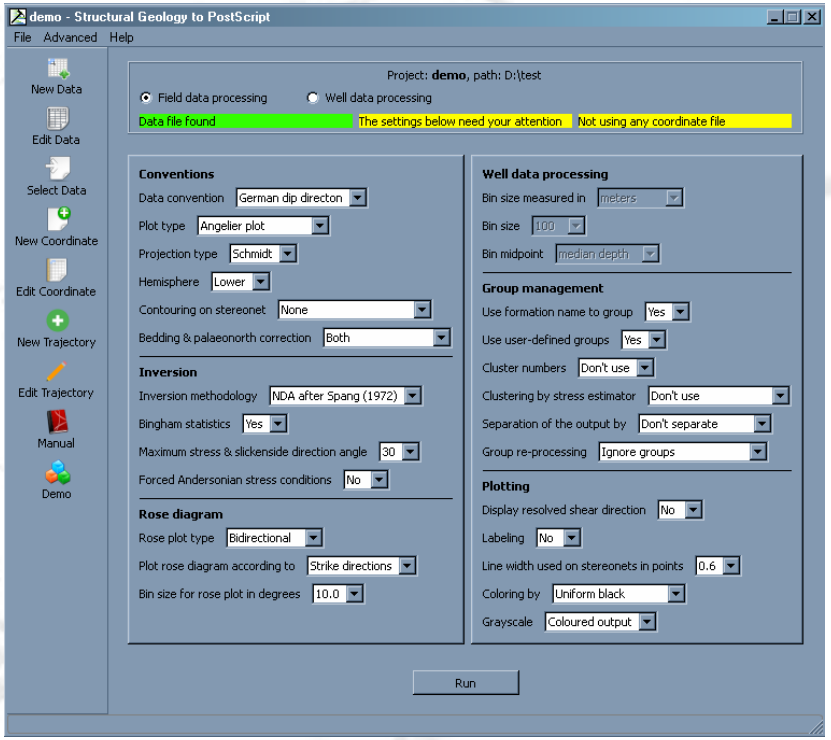
- Use '**NEW DATA**' to create a standard **input** (*.rgf) file
- Use '**EDIT DATA**' to edit your own input file with the default spreadsheet editor
- Use '**SELECT DATA**' to select input file to process

- Use '**NEW COORDINATE**' or '**EDIT COORDINATE**' to create/edit a standard **coordinate (*.xy)** file with the default spreadsheet editor; if a coordinate (*.xy) file is present that will be used, otherwise input file coordinates will processed
- Use '**NEW TRAJECTORY**' and '**EDIT TRAJECTORY**' to create/edit a standard coordinate (*.trj) file with the default spreadsheet editor; if a well data set is processed and a correct trajectory file is present, it will be used anyway

- The software generates the **project name** and searches for project **settings (*.set)**; if a correct one is found that will be processed, otherwise standard defaults will be displayed
- You can modify your settings in the '**SETTINGS**' field

- Push '**RUN**'

- Settings will automatically saved in the project folder in a **setting (*.set) file**
- **Processing result** and/or **working directory** with the project folder will displayed on the screen but will **NOT** be saved

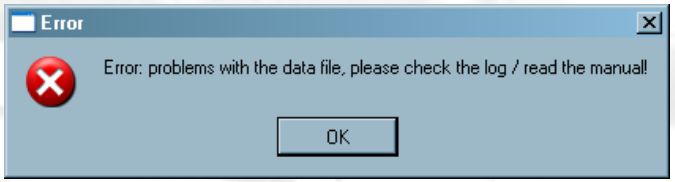


EDITING DATA FILES – MANUAL DISPLAY

- If a spreadsheet viewer and a PDF viewer is installed on your computer, the software will open to edit **input** (*rgf), **coordinate** (*xy) and **trajectory** (*.trj) file

DEMO

- Use this button first to test everything is OK with your **installation** and **pdf** viewer
- This option will ask for the **demo.rgf** file location, starts the **processing** of it, opens the **output folder**, generates the **output files** first in ***.eps** format and converts the into ***.pdf** later on. As the last step, the stereoplot output will be displayed in the **default pdf viewer** to test it

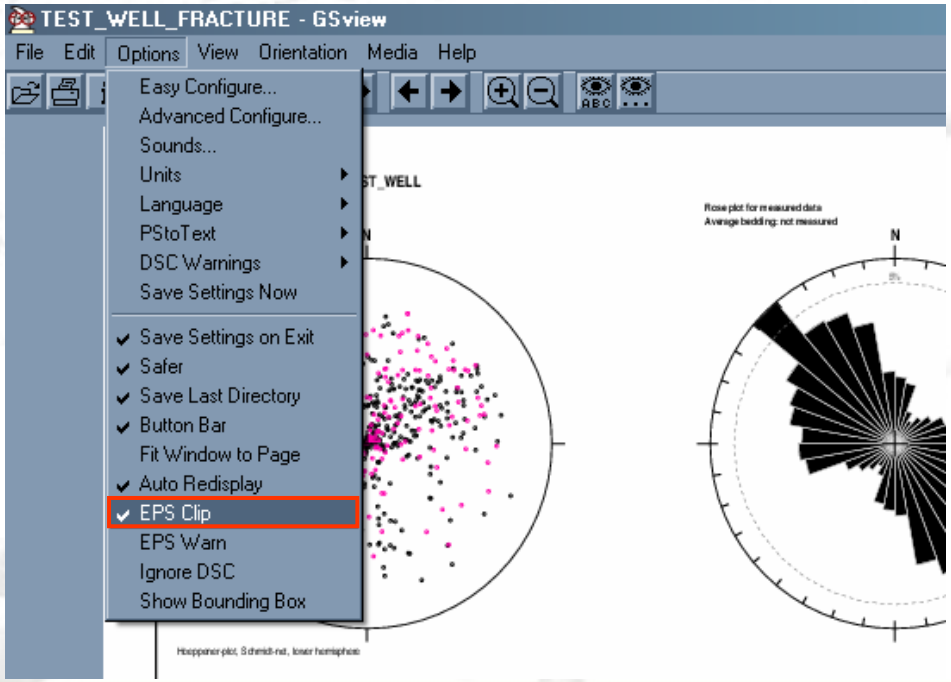


ERROR MESSAGE

- If the input (*.rgf), coordinate (*.xy) or trajectory (*.trj) file is incorrect, it will be not processed, and an error message will displayed

DISPLAY EPS RESULTS

- If you have **no pdf viewer**, or want to **see the *.eps file output**, you can use free software (e.g. **GSview**) to display generated *.eps files; if you want to use **GSview**, install **Ghostscript** before
- If you are using GSView, please don't forget to thick **Options/EPS clip option**, otherwise a rotated and clipped view will displayed on the screen
- You can convert the result vector graphs to ***.pdf** format and open it with a pdf viewer



PDF OUTPUT

- Conversion of the generated pdf's is available for the user's choice. To make the user's life easier, merged version of the individual pdf's is generated and dumped into the project directory.

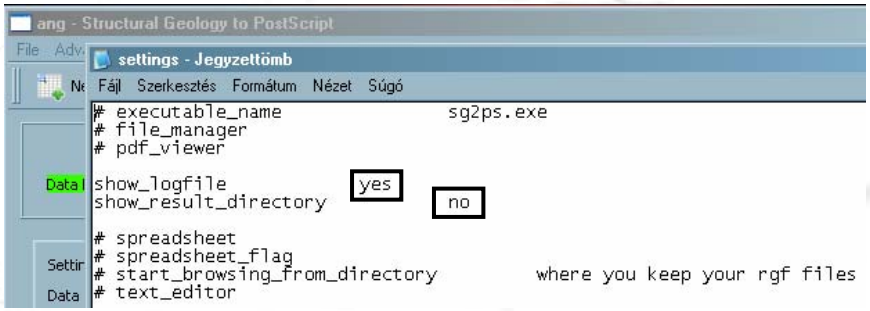
UNINSTALL

- Simply delete the application folder; no registry components were modified

INTRODUCTION

EDITING DATA FILES – MANUAL DISPLAY

- If you don't have default spreadsheet editor and/or PDF viewer, you can download **Apache OpenOffice** to process spreadsheet and **Ghostscript** for PDF display



ADVANCED MENU – GUI SETTINGS

- Here you can modify GUI settings
- Lines **with no # character** are active
- Use **yes/no** to display log file after processing, or to show working folder directory with the generated project folder
- After editing the settings, you **don't need** to restart the application

```
- For 'DO-01' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 244/75, s2: 140/04, s3: 049/15,     PURE EXTENSIVE, R: 0.681, R': 0.681, av. misfit: 19.0 deg.
- For 'DO-02' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 186/81, s2: 342/08, s3: 073/04,     PURE EXTENSIVE, R: 0.515, R': 0.515, av. misfit: 54.5 deg.
- For 'DO-03' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 227/31, s2: 014/55, s3: 127/16,   PURE STRIKE SLIP, R: 0.481, R': 1.519, av. misfit: 15.7 deg.
- For 'HA-05' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 183/09, s2: 080/54, s3: 279/34,      TRANSPRESSIVE, R: 0.181, R': 1.819, av. misfit: 44.3 deg.
- For 'KG-02' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 298/21, s2: 147/66, s3: 032/11,   PURE STRIKE SLIP, R: 0.514, R': 1.486, av. misfit: 11.2 deg.
- For 'KG-04' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 019/06, s2: 111/10, s3: 258/79, RADIAL COMPRESSIVE, R: 0.868, R': 2.868, av. misfit: 48.0 deg.
- For 'MA-01' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 241/67, s2: 335/02, s3: 066/23,       TRANSTENSIVE, R: 0.753, R': 0.753, av. misfit: 52.8 deg.
- For 'NE-01' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 156/84, s2: 318/06, s3: 048/02,     PURE EXTENSIVE, R: 0.457, R': 0.457, av. misfit: 52.3 deg.
- For 'NP-01' location, 'X' group (using group code), regression after Sprang (1972):
  - s1: 289/27, s2: 055/49, s3: 184/28,   PURE STRIKE SLIP, R: 0.492, R': 1.508, av. misfit: 58.9 deg.
```

PRE-PROCESSING

- Empty records of the input (*.rgf) file will be auto **completed**
- **Vectors** representing data input data will be generated

PROCESSING

- Using pitch data format, slickenside data will be **converted** into lineation format
- Using lineation data format for slickensides, **misfit correction** will be done (see Chapters 6.3.1 and 6.3.2)
- Slickenside **offset correction** is carried out (see Chapter 6.3.2);
- If it was required, **k-means clustering** will be carried out and the results will be displayed on the screen
- **Average data** for each data types (except lithology, slickenside and s-c data) is calculated for each location
- **Re-tilting** of all data with the average bedding will be done and slickenside offset re-correction will be carried out
- In the case of a well data set, trajectory correction and bedding correction is applied, and four data set (original, bedding corrected, trajectory corrected and a data set with both corrections) will be processed
- **Bingham statistics** and inversion results will be displayed for each fault or fracture data sets

FILE OUTPUT

- **Postscript** files, completed **geodatabase** files, **average** files and **data type by data type selected** files will be generated

- In the working folder, a **project folder** will be generated named after the current datum, time and project name
- Five output **subfolders** will be created in the project folder
- '**1_ORIGINAL**' folder contains the **copy of the original input** (*.rgf) file with no modifications as backup
- '**2_COMPLETED**' folder contains the **completed** version of the input (*.rgf) file with no empty records
- **3_AVERAGE** folder contains a file for the **average data** calculated for each different data types (except slickenside and s-c data) for each data points. This file is useful if you want to display average bedding in GIS
- '**4_RGF_SEPARATED**' folder has different subfolders containing input data **separated** by data points and data types
- The result **postscript** files are in '**5_PS_SEPARATED**' folder's subfolders, separated by data point, data types, groups and clusters (if this option was selected). No folder created if nothing to display
- '**6_WELL_PS_SEPARATED**' folder and its subfolders are for the well processing output, separated by wells, data type, groups and clusters for your choice. No folder created if nothing to display
- '**7_WELL_PLOTS_SEPARATED**' folder and its subfolders are for the well processing output in *.csv format as depth-value pairs, ready to load into any seismic interpretation application

Directory tree (right) generated after input processing ➡

```
20150920-063701_ST_WELL
├──1_ORIGINAL
├──2_COMPLETED
├──3_AVERAGE
├──4_RGF_SEPARATED
│   ├──BEDDING
│   └──FRACTURE
├──5_PS_SEPARATED
│   ├──BEDDING
│   └──FRACTURE
├──6_WELL_PS_SEPARATED
│   ├──BEDDING
│   └──FRACTURE
└──7_WELL_PLOTS_SEPARATED
    ├──BEDDING
    └──FRACTURE
```

Original (above) and completed (below) input file ⬇

DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	LDIP	SENSE	PALEONORTH	COMMENT
RW0508			RW226			LOWER BAKHTRIARI	bedding	162	7					
RW0509								171	20					
RW0510			RW228				bedding	228	8					
RW0511								194	10					
RW0512			RW230				bedding	165	18					
RW0513			RW232			QUATERNARY	LITHOLOGY							
RW0514			RW233				bedding	25	20					
RW0515							crossbedding	2	22					
RW0516			RW234				bedding	87	7					
RW0517								83	7					

DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	LDIP	SENSE	PALEONORTH	COMMENT
RW0508	X		RW226	36.1429	44.6694	LOWER BAKHTRIARI	BEDDING	162.0	7.0				0	
RW0509	X		RW226	36.1429	44.6694	LOWER BAKHTRIARI	BEDDING	171.0	20.0				0	
RW0510	X		RW228	36.1446	44.6724	LOWER BAKHTRIARI	BEDDING	228.0	8.0				0	
RW0511	X		RW228	36.1446	44.6724	LOWER BAKHTRIARI	BEDDING	194.0	10.0				0	
RW0512	X		RW230	36.145	44.6746	LOWER BAKHTRIARI	BEDDING	165.0	18.0				0	
RW0513	X		RW232	36.1449	44.6756	QUATERNARY	LITHOLOGY						0	
RW0514	X		RW233	36.1449	44.6756	QUATERNARY	BEDDING	25.0	20.0				0	
RW0515	X		RW233	36.1449	44.6756	QUATERNARY	CROSSBEDD	2.0	22.0				0	
RW0516	X		RW234	36.1433	44.6801	QUATERNARY	BEDDING	87.0	7.0				0	
RW0517	X		RW234	36.1433	44.6801	QUATERNARY	BEDDING	83.0	7.0				0	

```
============================
1) CHECKING OF SETTINGS FILE
============================

 - CHECKING V.SET SETTINGS FILE
   - Correct key/value pair for 'DATARULE:' user input key.
   - Correct key/value pair for 'PLOT:' user input key.
   - Correct key/value pair for 'PLOTTYPE:' user input key.
   - Correct key/value pair for 'HEMISPHERE:' user input key.
   - Correct key/value pair for 'CONTOURING:' user input key.
   - Correct key/value pair for 'TILTING:' user input key.
   - Correct key/value pair for 'GROUP:' user input key.
   - Correct key/value pair for 'CLUSTERNUMBER:' user input key.
   - Correct key/value pair for 'RUP_CLUSTERING:' user input key.
   - Correct key/value pair for 'GROUPSEPARATION:' user input key.
   - Correct key/value pair for 'INPUTGROUP:' user input key.
   - Correct key/value pair for 'INVERSION:' user input key.
   - Correct key/value pair for 'BINGHAM:' user input key.
   - Correct key/value pair for 'STRESSANGLE:' user input key.
   - Correct key/value pair for 'VIRTUAL:' user input key.
   - Correct key/value pair for 'ROSETYPE:' user input key.
   - Correct key/value pair for 'ROSEDIRECTION:' user input key.
   - Correct key/value pair for 'ROSEBINNING:' user input key.
   - Correct key/value pair for 'IDEALMOVEMENT:' user input key.
   - Correct key/value pair for 'LABELING:' user input key.
   - Correct key/value pair for 'LINEWIDTH:' user input key.
   - Correct key/value pair for 'COLORING:' user input key.
   - Correct key/value pair for 'GRAYSCALE:' user input key.
 - V.SET SETTINGS FILE IS CORRECT.

==============================
2) CHECKING OF INPUT DATA FILE
==============================

 - CHECKING V.RGF INPUT DATA FILE
    - The header (column names) of the data file has been processed.
    - The following columns are reserved but unused:
       "DEPTH"
    - Input V.RGF file read, 2864 record(s) imported.
    - Existing DATA_ID's in all records.
    - Correct DATA_ID's in all records.
    - Correct group code(s) in all records.
    - Correct color code(s) in all records.
    - Correct coordinate(s) in all records.
    - Correct datatype(s) in all records.
    - ERROR: incorrect strike/dip direction(s) in the following record(s):  BE-03-11.
    - Correct dip angle(s) in all records.
    - Correct CS/striae(s) in all records.
    - Correct paleo north direction(s) in all records.

 WARNING, RGF_ERROR: the input V.RGF file structure is incorrect, the file will not be processed.
```

PRE-PROCESSING

- Before processing, input (*.rgf) file will be **checked**
- If the file contains error(s), **type** of error and the incorrect **record ID** (from input file) will displayed
- In the case of any error, 1) **open** your input file with text editor application, 2) **fix** the incorrect record, 3) **save** the input file and process it again
- Because input database error check is carried out column by column, a need to **repeat** error fixing for other columns can easily arise

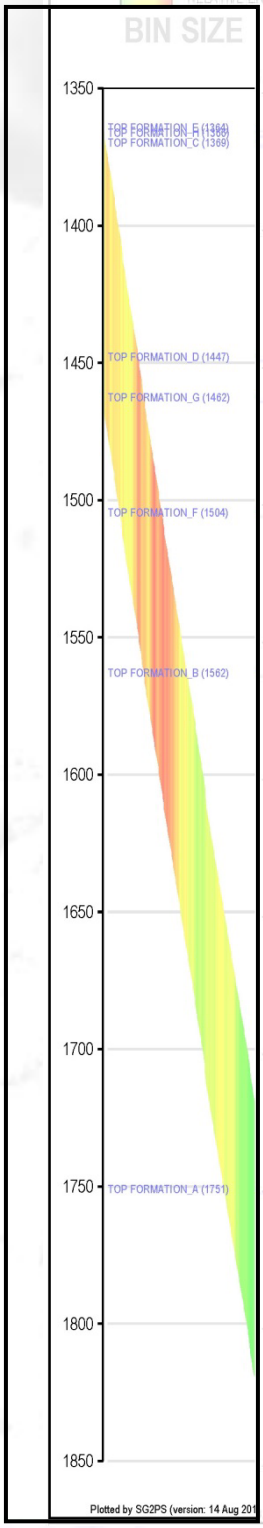
PROCESSING

- If data number is **less** than required by the slickenside inversion procedure or Bingham statistics, the software is unable to calculate stress field or statistics parameters, and an error message will displayed, but the program will continue the processing
- If regression is required for the average calculation and the data set is **singular**, no average will be calculated, and a message will inform the user about the data set singularity

WELL AND FIELD DATA GRAPHICAL OUTPUT

WELL DATA GRAPHICAL OUTPUT

BIN SIZE PANEL

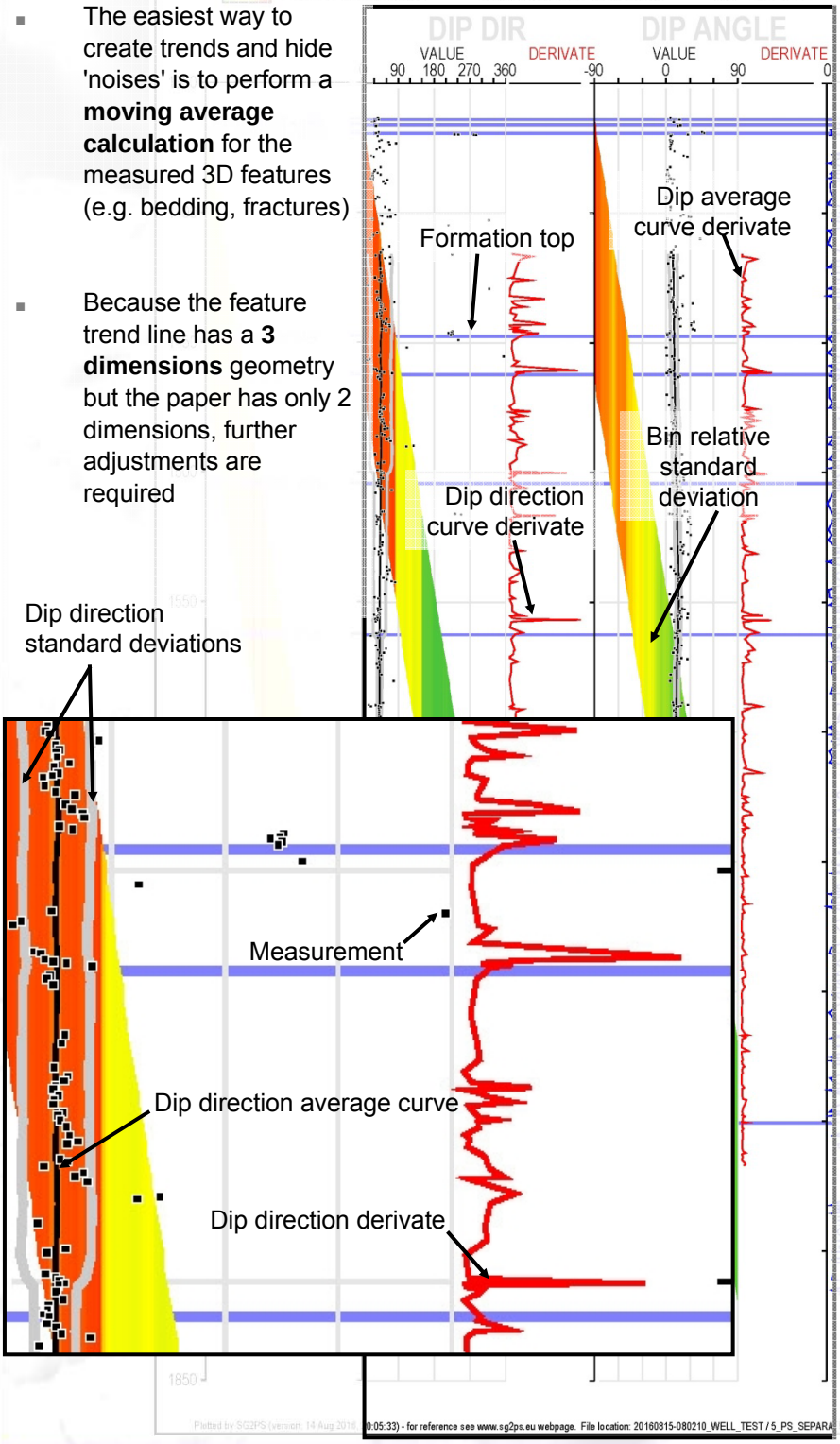


- The main philosophy behind the well data structural processing is to try to identify and visualize the **main deformation zones**
- The preliminary assumption is on the relationship between the measured features geometry and the amount of deformation: if a significant and **abrupt change** is visible in the features geometry trend, it could be an indication of a **deformation zone**
- In order to find significant and abrupt changes in the features geometry trend, the **trend** itself has to visualize first
- The easiest way to create trends and hide 'noises' is to perform a **moving average** calculation for the measured 3D features (e.g. bedding, fractures)
- In order to do a moving average calculation, a **bin size** has to be defined first (see Chapter 5.1). The bin represents an **interval for data processing**, and can be a **fix length** or a **fix data number** interval. These intervals are displayed here by colored vertical bars
- It is important to have feedbacks on the reliability of the moving average calculation. One of the possible feedbacks is the **relative data number** in each bins (if fix interval was used)
- If a bin has relatively **large number** of data (displayed by green), then the moving average calculation should be **reliable**.
- On the other hand, a bin with relatively **small number** of data (marked by red) is **not reliable**
- Please note this visualization is not performed if fix data number bin is used

WELL DATA GRAPHICAL OUTPUT

DIP DIRECTION AND DIP ANGLE PANEL

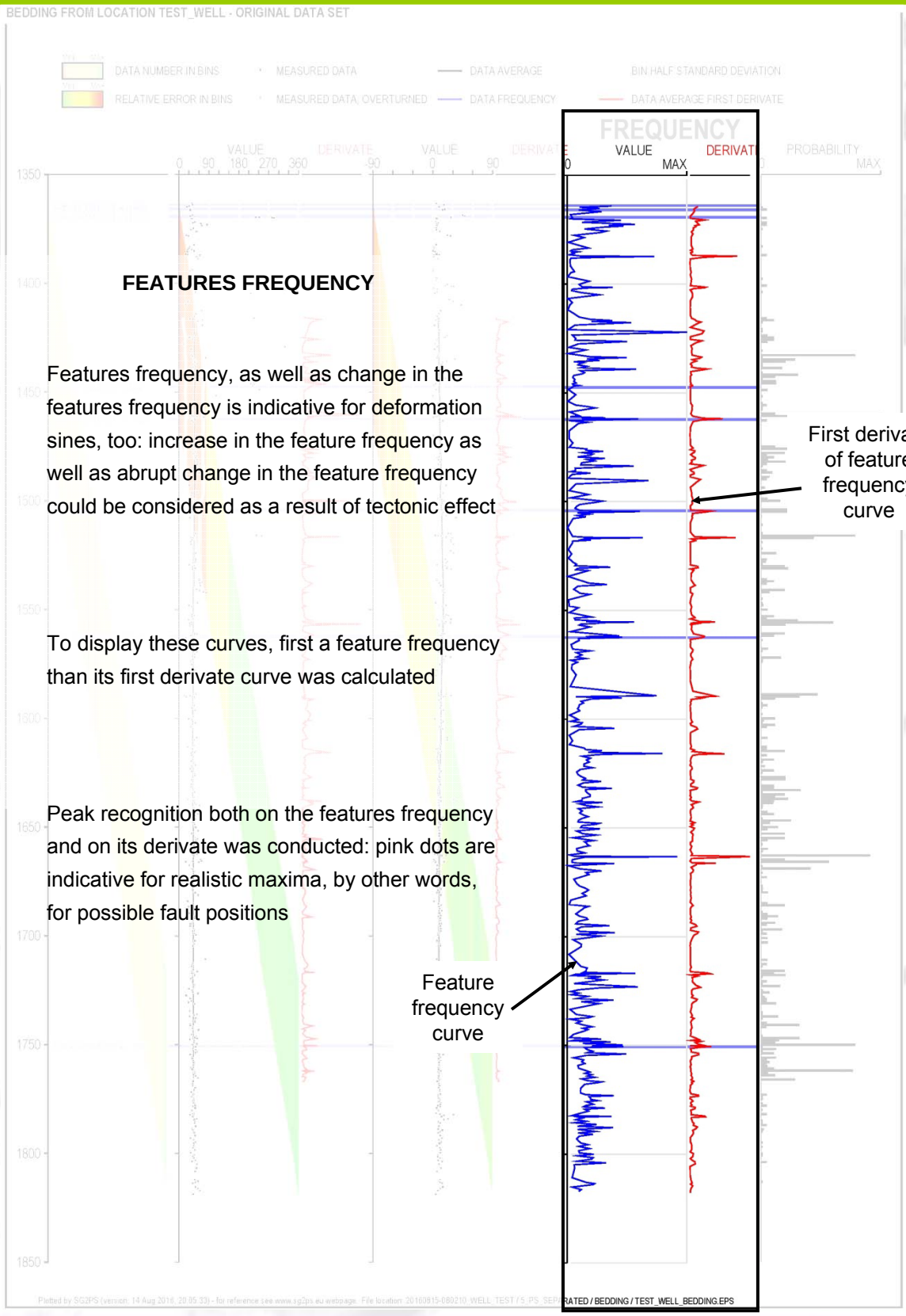
- The easiest way to create trends and hide 'noises' is to perform a **moving average calculation** for the measured 3D features (e.g. bedding, fractures)
- Because the feature trend line has a **3 dimensions** geometry but the paper has only 2 dimensions, further adjustments are required



- To follow classic geological data handling, the average 3D feature trend line is **'decomposed'** into two lines and displayed on two panels: one for the **dip direction** and one for the **dip angle**
- **Abrupt changes** (=positive or negative local maxima) of the average feature trend line could be indicative for **structural effects**
- It is important to have feedbacks on the **reliability** of the moving average calculation. The method is to calculate the **standard deviation curve** for the dip direction and dip angle in each data bins, and display it together with the trend line curve
- Here a **relative standard deviation** is displayed by colored bars for each bins, too: if a bin has relatively large error (displayed by red), then the moving average calculation should not be reliable. On the other hand, a bin with relatively small error (marked by red) could be reliable
- To visualize **how intensive** the change in the structural trend, **first derivate** of both the dip direction the dip average trend line is displayed
- Abrupt changes (=**positive or negative local maxima**) of the first derivate of the average features trend line could be indicative for **structural effects**
- Individual **measurements** are indicated on the same 'decomposed' way: measurement dip direction on the first panel, and measurement dip on the second one
- If **less data than required** by the processing (pro example user has a 100m window, but the well section is shorter), then the measurements will be displayed with no average, standard deviation and derivate curves are plotted

WELL DATA GRAPHICAL OUTPUT

FEATURE FREQUENCY PANEL

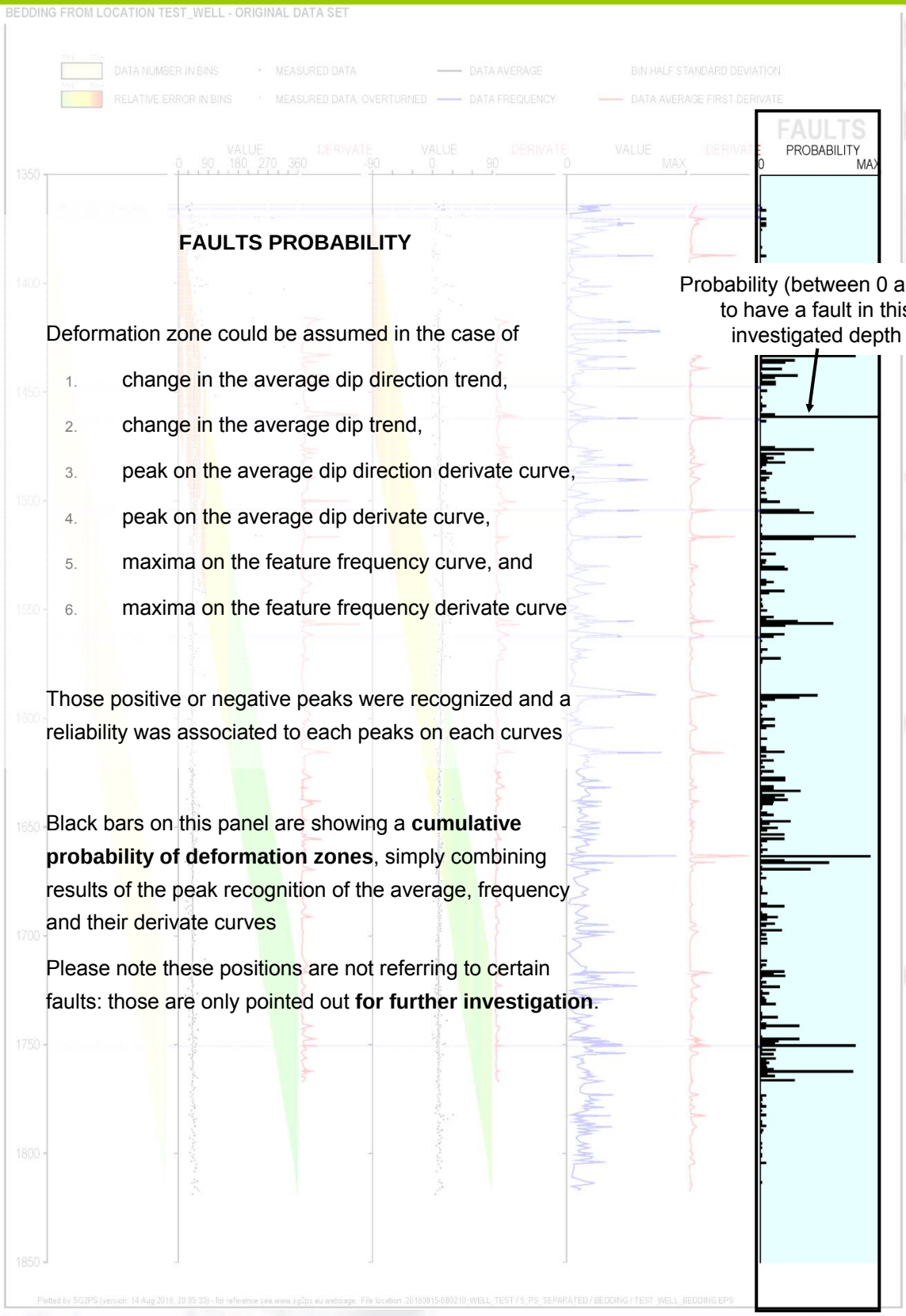


FEATURES FREQUENCY

- Features frequency, as well as change in the features frequency is indicative for deformation sines, too: increase in the feature frequency as well as abrupt change in the feature frequency could be considered as a result of tectonic effect
- To display these curves, first a feature frequency than its first derivate curve was calculated
- Peak recognition both on the features frequency and on its derivate was conducted: pink dots are indicative for realistic maxima, by other words, for possible fault positions

WELL DATA GRAPHICAL OUTPUT

FAULTS PROBABILITY PANEL



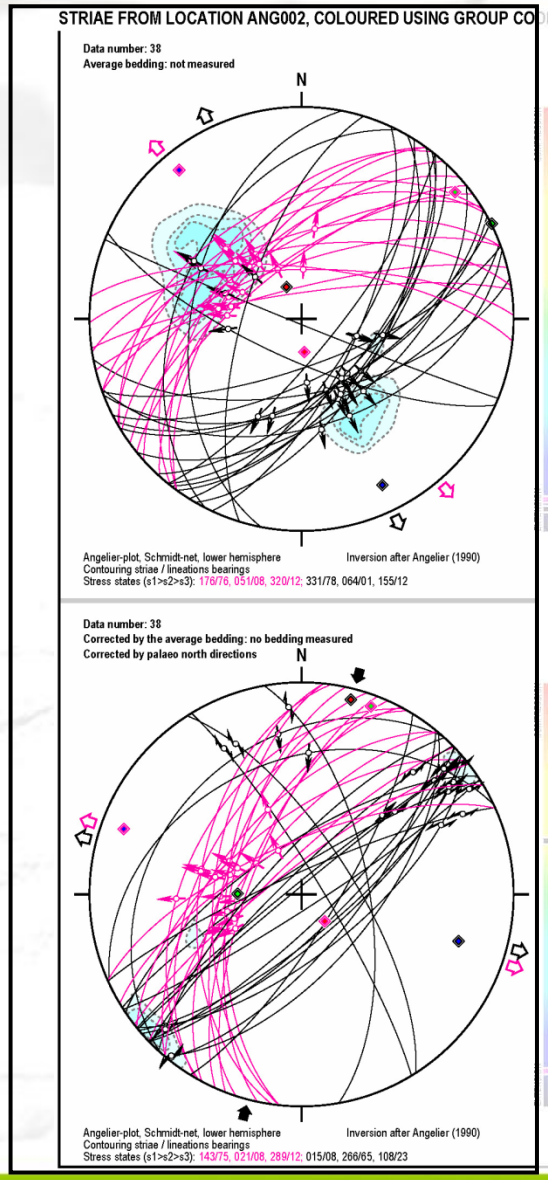
FAULTS PROBABILITY

Probability (between 0 and 1) to have a fault in this investigated depth

- Deformation zone could be assumed in the case of
  1. change in the average dip direction trend,
  2. change in the average dip trend,
  3. peak on the average dip direction derivate curve,
  4. peak on the average dip derivate curve,
  5. maxima on the feature frequency curve, and
  6. maxima on the feature frequency derivate curve
- Those positive or negative peaks were recognized and a reliability was associated to each peaks on each curves
- Black bars on this panel are showing a **cumulative probability of deformation zones**, simply combining results of the peak recognition of the average, frequency and their derivate curves
- Please note these positions are not referring to certain faults: those are only pointed out **for further investigation**.

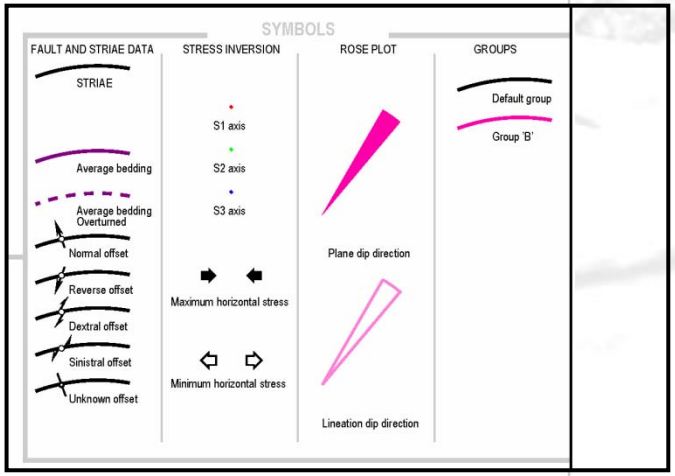
FIELD DATA GRAPHIC OUTPUT

STEREOGRAPHICAL PROJECTIONS



- You can use free softwares (e.g. **GSview**) to display generated *.eps files; if you want to use **GSview**, install **Ghostscript** before
- If you are using GSView, please don't forget to thick **Options/EPS clip option**, otherwise a rotated and clipped view will displayed on the screen
- At the left upper corner, **data type, data location** is indicated
- If the option 'using groups' is selected, than **group name** is indicated there
- If used, the **formation name** is indicated here
- Upper half of the final plot is to show **original** data set, lower half is for the **re-tilted** (bedding and/or palaeonorth corrected) visualization of them
- Plot **type, projection and hemisphere**, as well as **result of stress inversion** is presented below the stereonets

- If fractures were processed, Bingham analysis could be displayed here:
  - density ellipsoid **axis directions** and
  - **axes length** (eigenvalues)
- For slickensides, stress inversion is carried out, red, green and blue rectangles show the **maximum, intermediate and minimum stress axis** orientation



- For planes, Bingham statistics could be used; red, green and blue rectangles show the **maximum, intermediate and minimum weight axis** orientation
- Large arrows show the **maximum and minimum horizontal stress** directions
- Dark purple great circle shows the **average bedding**

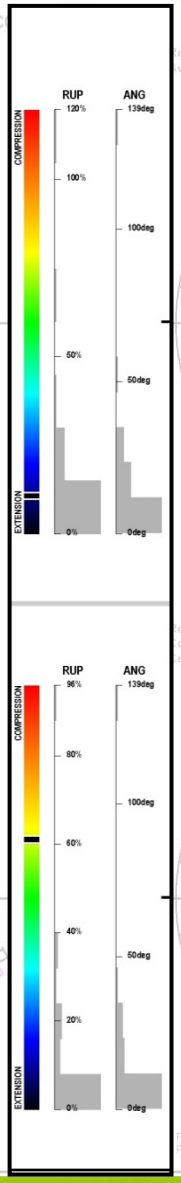
FIELD DATA GRAPHICAL OUTPUT

STRESS FIELD PROPERTIES, STRESS ESTIMATORS

- Rainbow colored bar shows the possible stress regimes from radial extensive (black) to radial compressive (red). White line shows the **current stress regime** after inversion

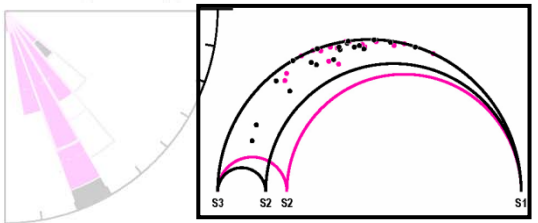
- **RUP** bar shows the **relative upsilon ratio** (theoretical / resolved shear stress vector length) from 0 to maximum RUP value (see Chapter 6.9.2). Thickness of gray bars in the background shows the density of RUP values

- If different clusters are visible on the **RUP** chart, they can show different striae sets working in different stress fields; use RUP bar to **identify possible striae sets** in a mixed data set



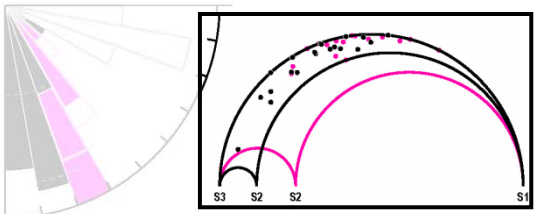
- **ANG** bar shows the **angular misfit** between the measured and calculated stress directions from 0 to the maximum ANG value. Thickness of gray bars in the background shows the density of ANG values

- Use **ANG** bar to check **reliability** of stress tensor calculation; large number of large misfits indicates uncertain / questionable stress inversion result



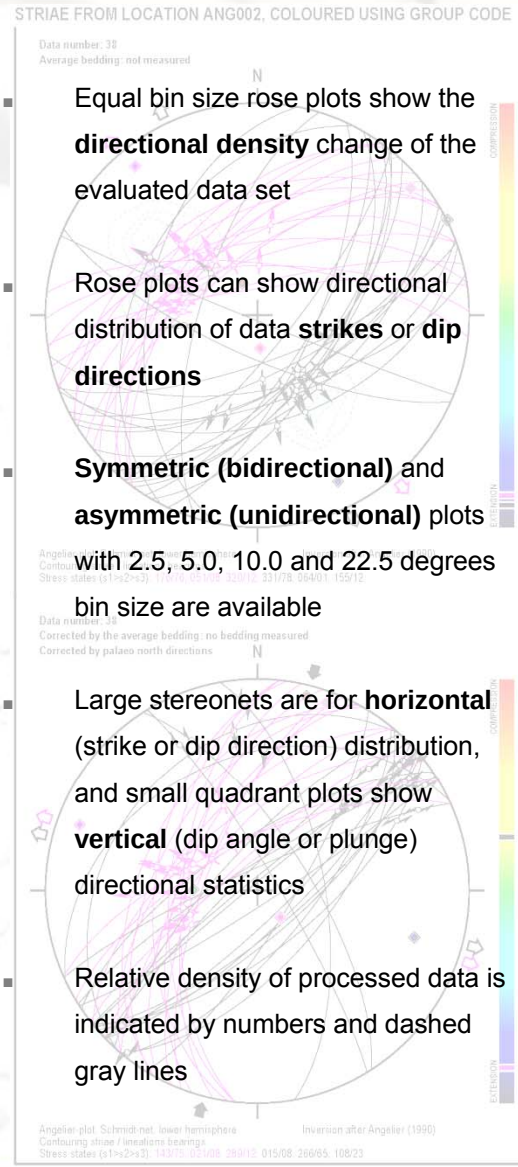
- **Mohr circles** show the **stress status** as well as normalized normal and shear stress values for each evaluated data

- If more than one group displayed, then **color** of Mohr circles refers to the relevant group

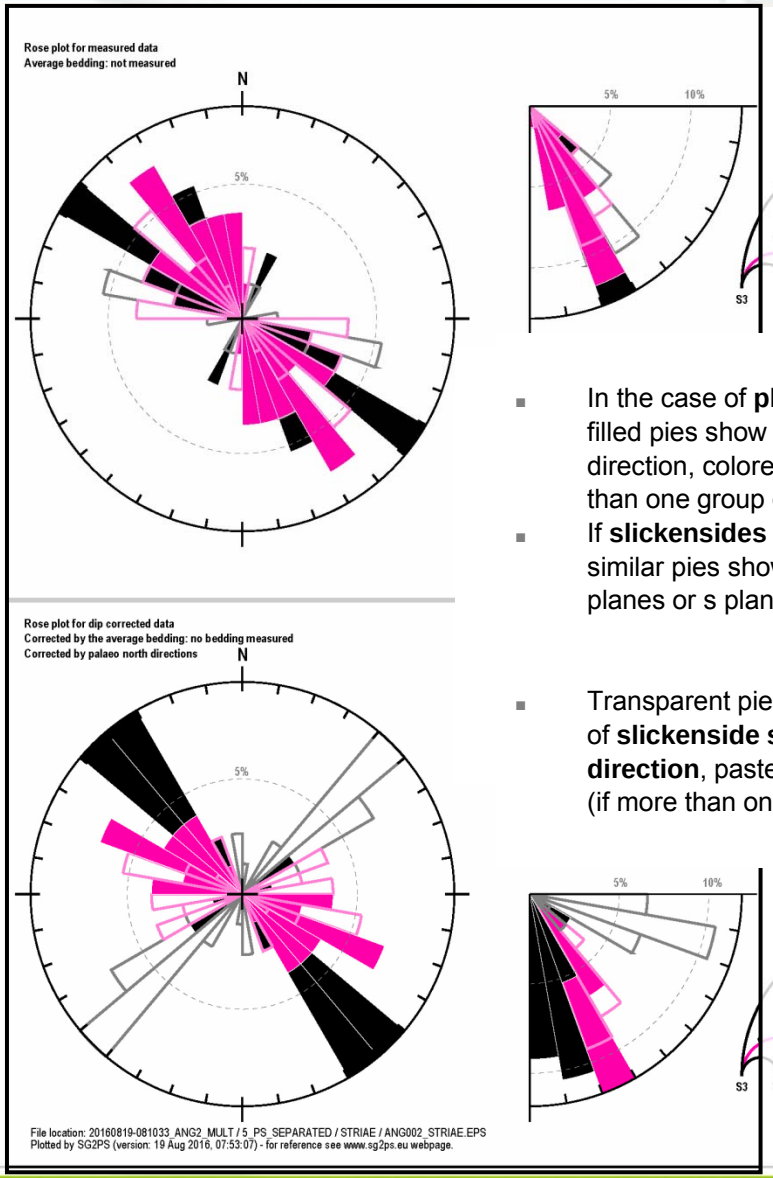


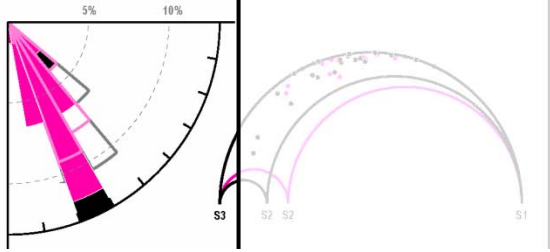
2.2.3 FIELD DATA GRAPHIC OUTPUT

ROSE PLOTS



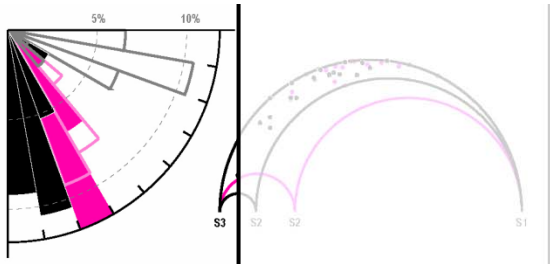
- Equal bin size rose plots show the **directional density** change of the evaluated data set
- Rose plots can show directional distribution of data **strikes** or **dip directions**
- **Symmetric (bidirectional)** and **asymmetric (unidirectional)** plots with 2.5, 5.0, 10.0 and 22.5 degrees bin size are available
- Large stereonets are for **horizontal** (strike or dip direction) distribution, and small quadrant plots show **vertical** (dip angle or plunge) directional statistics
- Relative density of processed data is indicated by numbers and dashed gray lines





- In the case of **planes and lineations**, color filled pies show relative density of data dip direction, colored by the relevant group (if more than one group displayed)
- If **slickensides or s-c planes** are evaluated, similar pies show relative density of striated planes or s planes

- Transparent pies are displaying relative density of **slickenside slip direction** or **c plane dip direction**, pastel colored by the relevant group (if more than one group displayed)



GROUPS
AND
COLOR MANAGEMENT

GROUPS AND COLOR MANAGEMENT

IMPORT AND MAKE GROUPS

- Each user-created input file could have a user defined **1-digit group code**, in the range of A – I. If any record has no group code, will get a default 'X' character. The user can choose use these group codes during the processing or not

⬇

- While k-means clustering, a second character will be added in the range of 'A' – 'I', referring to each clusters, resulting in **a 2-digits group code**. If no clustering, a default 'X' character will be added each data set

⬇

- While RUP/ANG clustering of slickensides, a third character will be added in the range of 'A' – 'I', referring to each clusters, resulting in **a 3-digits group code**. If no clustering or no striae to process, a default 'X' character will be added each data set

⬇

- Each output file will have a **3-digits group code**: the first character will be the **user-define one**, the second one shows the result of the **k-means clustering**, and the third one is for the **RUP/ANG clustering of slickensides**

ORIGINAL FILE

⬇

A__

⬇

K-MEANS CLUSTERING

⬇

AB_

⬇

RUP/ANG CLUSTERING

⬇

ABX

⬇

OUTPUT FILE

- You can process your groups, and in the next step while exporting, you can export and color them on many different ways, combining the options below:

1) PROCESSING OF GROUPS

While processing, you have the following options

1. **Ignore** the **user defined 1-digit group code**, and process everything together
2. Use the **user defined 1-digit group code**, and process everything group by group

2) EXPORTING OF GROUPS

While exporting, you have the following options

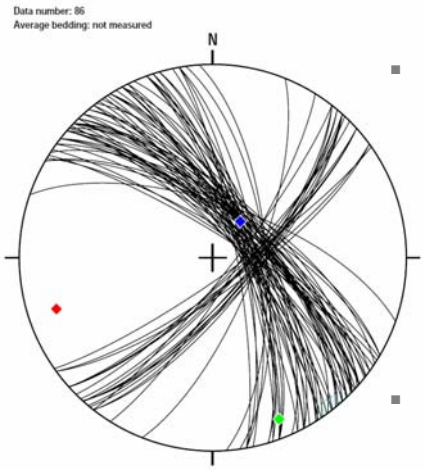
1. **Ignore** your groups, and put everything into the same output file
2. Use the **first** character (user defined group), to separate your data by the original user defined groups
3. Use the **second** character (k-means clustering result), separate your data by the k-means clustering result
4. Use the **third** character (RUP/ANG clustering result), separate your data by the RUP/ANG clustering result

3) COLORING OF GROUPS

While coloring, you have the following options

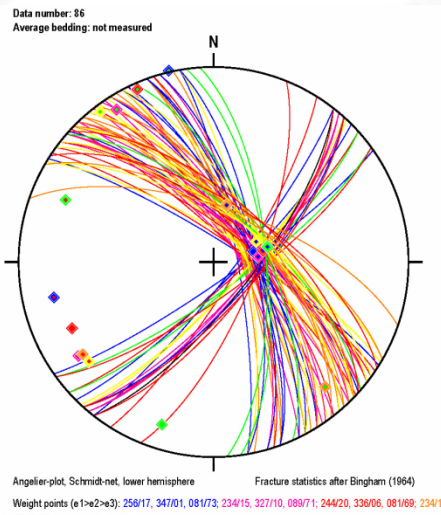
1. **Ignore** any groups and color code, and color everything black
2. Use the **color code** to color everything
3. Use the **first** character of the group code (user defined group), to separate your data by the original user defined groups
4. Use the **second** character of the group code (k-means clustering result), separate your data by the k-means clustering result
5. Use the **third** character of the group code (RUP/ANG clustering result), separate your data by the RUP/ANG clustering result

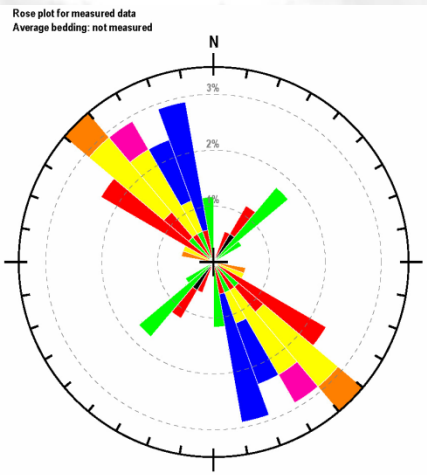
MULTIPLE GROUPS ON THE SAME CHART



It is possible to separate the processed data set 1) by user defined group codes, 2) by k-means clustering result, 3) by RUP/ANG clustering results (in the case of slickensides), or 4) display everything together

To find coloring options, please see Chapter 2 and 5.21

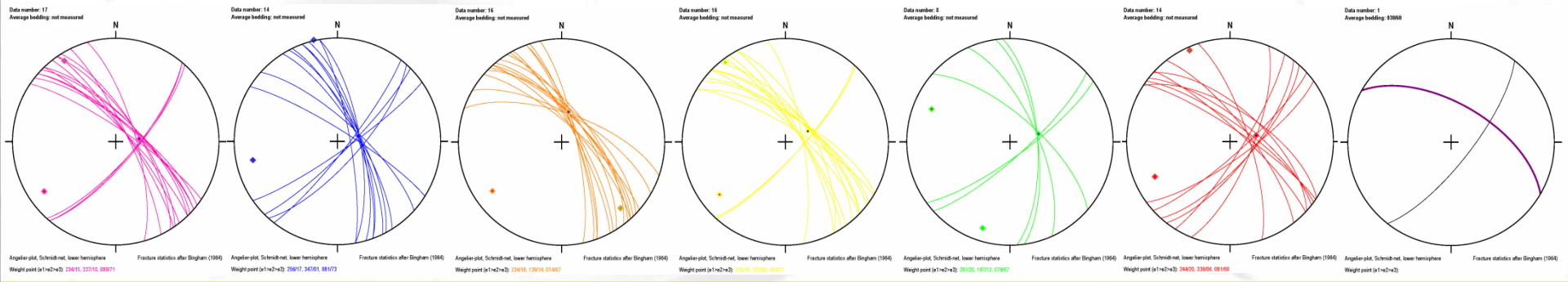




⬆ Original data set, ignoring both user defined color codes and groups

⬆ Exporting a data set **without** separation by user define groups, showing everything (used defined groups A, B, C, D, E, F and X) on the same stereonet. Data is colored by data sets group code

Exporting the same data set **separated by their user define groups**, resulting in 7 stereonets, one for each group code (A, B, C, D, E, F, X). Data is colored by data sets group code ⬇



INPUT FILE FORMAT

4.1.1 INPUT FILE FORMAT

RGF DATA FORMAT I.

- RGF data file (*.rgf) is a text file. **First line is the header**, with **FIXED** content (used by the software to identify expected field content) to indicate column names; header column names are **bold and blue** below. Columns could be in arbitrary order; header is **case sensitive**, the file content is not. Please use ANSI characters for the file name and content
- 1st column is for the **DATA_ID**; all of measurements entered in the RGF file **must** have an individual ID, which **must be different** from each other
- 2nd column shows the **GROUP_CODE**; use if you want to process more than one group (ie. different slickenside sets on the same location); characters 'A' – 'I', 'X', 'a'-'i' and 'x' are accepted here. If the field is empty, group will became default group 'X'. Use one digit for an original input file; the output will get a three digits GROUP_CODE, first character for the user defined groups, second one for the k-means clustering, and a third one for the RUP/ANG clustering results. If no e.g. RUP/ANG clustering, the third digit will became 'X'
- 3rd column is for the **COLOR_CODE**; use characters 'A' – 'J', 'a' – 'j' or numbers 0 – 9 to plot different colors. Empty cell, A and 0 is for default black color, other characters are for the colors at the right
- 4th column is for the **LOCATION** of the data point where the measurement has done. You just need to enter the **new** data point name, otherwise the data point name of the previous record will be used. At least the first record must contain data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	L_DIP	SENSE	PALEONORTH	DEPTH	COMMENT
2	P_S_0019	a	e	547	1436887	366576	Dachstein	Striae	258	78	188	58	D			
3	P_S_0020	a	e						271	73	198	N	D			reliable
4	P_S_0021	a	e					Bedding	264	64						overturned
5	P_S_0022	a	e						278	77						
6	P_S_0029	a	f						266	76						
7	P_S_0030	a	f	Well-2	1436656	355889	Dachstein	Bedding	264	74				12	1345.55	
8	P_S_0031	b	f				Kossen	Fracture	248	69				12	1445.76	
9	P_S_0032	b	f						272	72				12	1657.99	

- 5th and 6th columns are for the **LOC_X** and **LOC_Y** **coordinates** of the data point; just enter **new** data, otherwise previous record will be used. The following solutions are useful:
  - leave these fields empty if you do not have coordinates or do not need them
  - fill them manually, or
  - if you want to process the final spread sheet with GIS / database manager softwares, use **coordinate** (*.xy) file (see Chapter 4.3) – in this case the content of these columns will be **overwritten** by the content of xy coordinate file
- 7th column is for the **FORMATION** name; not required to fill. Enter new formation name, otherwise it will be filled by content of previous record
- 8th column is for the **DATATYPE**. Enter new data type if changed, otherwise it will be filled by content of previous record:
  - for **lithology**: 'lithology'
  - for **lineations**: 'boudain', 'foldaxis', 'kink', 'lineation', 'userlineation1', 'userlineation2', 'userlineation3', 'userlineation4' and 'userlineation5'
  - for **planes**: 'bedding', 'borehole break-out' or 'bbo', 'contact', 'crossbedding', 'foldplane', 'foldsurface', 'fracture', 'lithoclase', 'plane', 'userplane1', 'userplane2', 'userplane3', 'userplane4', 'userplane5', 's1', 's2', 's3', 's4', 's5', 'stress-induced fracture' or 'sif', and 'vein'
  - for **slickensides**: 'striae', and
  - for **s-c schistosity**: 'sc'

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	L_DIP	SENSE	PALEONORTH	DEPTH	COMMENT
2	P_S_0019	a	e	547	1436887	366576	Dachstein	Striae	258	78	188	58	D			
3	P_S_0020	a	e						271	73	198	N	D			reliable
4	P_S_0021	a	e					Bedding	264	64						overturned
5	P_S_0022	a	e						278	77						
6	P_S_0029	a	f						266	76						
7	P_S_0030	a	f	Well-2	1436656	355889	Dachstein	Bedding	264	74				12	1345.55	
8	P_S_0031	b	f				Kossen	Fracture	248	69				12	1445.76	
9	P_S_0032	b	f						272	72				12	1657.99	

- 9th **DIP_DIR** column is for the **strike/dip direction** (depending on input settings) of lineations, planes, planes with slickenside, or s planes of s-c data sets. All of records must be filled except the case of 'lithology' datatype. Values between 0.0 – 360.0 are accepted

- 10th column is for the **DIP angle** (plunge) of lineations, planes, planes with slickensides, or s planes of s-c data sets. All of records must be filled except the case of 'lithology' datatype with values between 0.0 – 90.0

- 11th **L_DIP_DIR** column is
  - for the **strike/dip direction** either **c planes** of s-c data sets, or for the bearing of **striae lineation** (values between 0.0 – 360.0 are accepted), or
  - for the **pitch angle** of the striation in the case of using pitch convention. All of records must be filled in the case of 'striae' or 'sc' data type; enter values between 0.0 – 90.0 for the pitch angle
- 12th **L_DIP** column is
  - for the **dip angle** of **c plane** or plunge of **slickenside lineation**. All records must filled if using 'striae' or 'sc' data types (values between 0.0 and 90.0 are accepted), or
  - in case of **pitch** convention this column indicates the **dip direction of the pitch** (enter 'n', 'nne', 'ne', 'ene', 'e', 'ese', 'se', 'sse', 's', 'ssw', 'sw', 'wsw', 'w', 'wnw', 'nw' or 'nnw' for pitch direction)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	L_DIP	SENSE	PALEONORTH	DEPTH	COMMENT
2	P_S_0019	a	e	547	1436887	366576	Dachstein	Striae	258	78	188	58	D			
3	P_S_0020	a	e						271	73	198	N	D			reliable
4	P_S_0021	a	e					Bedding	264	64						overturned
5	P_S_0022	a	e						278	77						
6	P_S_0029	a	f						266	76						
7	P_S_0030	a	f	Well-2	1436656	355889	Dachstein	Bedding	264	74				12	1345.55	
8	P_S_0031	b	f				Kossen	Fracture	248	69				12	1445.76	
9	P_S_0032	b	f						272	72				12	1657.99	

- 13th **SENSE** column is for the 1) **movement direction** along slickenside, and 2) to indicate **normal/overturned** geometry of bedding. Must be filled in the case of slickensides and in the case of overturned beds, otherwise it is empty. The following characters are accepted here for slickensides:
  - '+', 'thrust', 'up', 'inverse', 'u' or 'i' for **inverse/reverse** offset
  - '-', 'normal', 'fault', 'down', 'downward', or 'n' for **normal** movement
  - 'dextral', 'dx' or 'd' for **dextral** movement
  - 'sinistral', 'sn' or 's' for **sinistral** movement, and
  - 'x' or 'none' for **unknown** offset

  For bedding planes, use
  - 'o' or 'overturned' to indicate **overturned** bedding, otherwise it will be handled as normal one
  - it is allowed (but not necessary) to use 'n' or 'normal' to indicate **normal** bedding
- 14th column is for the **PALEONORTH direction**; if it is known, can be indicated and used, otherwise leave empty. Values between 0.0 – 360.0 are accepted here
- 15th column is for the well data **DEPTH** which must be present for each well data sets
- 16th column is for your **COMMENTS**
- 17th and other columns: will be not processed by the software. Everything **behind** the 17th column will be exported with no modification, so it is possible to store here as many data as you want

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DATA_ID	GROUP_CODE	COLOR_CODE	LOCATION	LOC_X	LOC_Y	FORMATION	DATATYPE	DIP_DIR	DIP	L_DIP_DIR	L_DIP	SENSE	PALEONORTH	DEPTH	COMMENT
2	P_S_0019	a	e	547	1436887	366576	Dachstein	Striae	258	78	188	58	D			
3	P_S_0020	a	e						271	73	198	N	D			reliable
4	P_S_0021	a	e					Bedding	264	64						overturned
5	P_S_0022	a	e						278	77						
6	P_S_0029	a	f						266	76						
7	P_S_0030	a	f	Well-2	1436656	355889	Dachstein	Bedding	264	74				12	1345.55	
8	P_S_0031	b	f				Kossen	Fracture	248	69				12	1445.76	
9	P_S_0032	b	f						272	72				12	1657.99	

4.2.1 INPUT FILE FORMAT

SET DATA FORMAT I.

SET data file (*.set) is a **non case sensitive**, **tab separated** text file storing the settings **keys** and **values** used during data processing. It is not necessary to have all settings; if any correct setting is found in the file will be used, otherwise default settings were applied. The order of these settings in the file is arbitrary. Default keys are **bold and blue**, and default settings values (hard coded ones in software) are **bold** characters

0) PROCESSING MODE

WELLDATA: '**n**' to evaluate input file as a field data set, and 'y' to process the input file as a well data set

1) WELL

WELLINTERVAL (5.1) while processing well data (e.g. fractures or bedding), a moving average of the processed data set is calculated. The interval, used for the step-by-step moving average calculation is called *bin*. Size of this bin can be either defined in **meters** and you can use '**m**', or if you like, you can create fix data number interval applying 'd' in the settings file

WELLINTERVAL_LENGTH (5.2) here you can define the length of your bin: how many meters or how many piece of data can belong to one bin between 1 and 5000 (**100** meters or piece of data as default). We recommend to have at least 50-100 bins, so divide the complete data number or your logged interval by 50-100

WELLINTERVAL_MIDDLE (5.3) here the question is how to calculate the middle depth of the processed bin: it can be either the median depth of each interval (use '**m**'), or the arithmetic average (use 'a') of each intervals

2) PLOT

DATARULE (5.4): '**g**' for German dip direction/dip convention, and 'r' for right hand rule input (output will be in the same format as the input)

PLOT (5.5): '**a**' for classic Angelier plot, and 'h' for Hoeppner pole point plot

PLOTTYPE (5.6): '**s**' for equal area Schmidt projection, and 'w' for equal angle Wulff projection

HEMISPHERE (5.7): 'u' in the case of upper, and '**l**' for lower hemisphere

CONTOURING (5.8): enter '**n**' do not use, 'd' to contour dip direction/lineation bearing, 's' for strike/lineation bearing, 'o' for plane normal/lineation bearing and 'b' for slickenside/lineation bearing

TILTING (5.9): refers to the **bedding and palaeonorth direction correction**: use 'b' to re-tilt everything with the bedding, 'p' to correct by palaeonorth direction, and '**a**' to do both of them

3) INVERSION

INVERSION (5.10): for the striae **inversion** methodology:

- 'n' for no inversion,
- 'b' for brute force inversion,
- '**d**' for NDA after Spang (1972),
- 'p' for PTN after Turner (1953),
- 'f' for Fry (1999),
- 's' for Shan et al (2003),
- 'm' for Michael (1984),
- 'a' for Angelier (1990), and
- 'o' for Mostafa (2005)

BINGHAM (5.11): for fracture **weight point statistics**: enter '**b**' to use Bingham statistics, or 'n' if you do not need

STRESSANGLE (5.12): for the **angle** between the maximum stress direction and the slickenside direction, and will be used only by NDA and PTN methodologies; enter a number between 10 and 80 (**30** deg is the default setting)

VIRTUAL (5.13): for the **virtual symmetrical data set**: enter 'y' to use them, otherwise input '**n**'

4) DATA GROUP MANAGEMENT

INPUTGROUP (4.14): this option is used while opening the result of a previous interpretation loop. The output file has a three digits group code (1st built up by the user defined one, 2nd is the k-means clustering result and 3rd for the RUP/ANG clustering). You can ignore them ('**n**'), use the first digit ('f') to group your data for the original user defined group, or use the second digit ('s') to group your data using the former the k-means clustering results,(if it exists at all), or use the third digit ('t') to group your data using the former RUP/ANG clustering clustering results

GROUP (3, 4.15): if you want to use user defined groups (if they are indicated in 2nd column of RGF file), all data will be separated by locality, by datatype **plus by your groups**, and processed/displayed on this way. Input '**y**' to use groups, or 'n' to do not use them

FORMATION: enter 'n' to do not separate your data set by formations, and use 'y' if you want to do so

CLUSTERNUMBER (4.16): enter '**n**' do not use, 'a' to find the ideal cluster number, or 2 – 9 for the required number of clusters. Either the entire data set will be processed, or clustering will be done group by group

RUP_CLUSTERING (4.17): enter '**n**' do not use, 'a' to do a clustering using the angular misfit (ANG) values of the stress inversion, or enter 'r' to use the relative upsilon (RUP) value (see Chapter 6.9.2) in the clustering procedure

GROUPSEPARATION (2): you had the chance to group everything by a user defined group code, by k-means clustering result and later by RUP/ANG clustering. While creating the output files, group codes can be ignored ('**i**') and each groups will handled together, or your data set will be displayed on different stereonets, separated by the user defined group code ('g'), by the k-means clustering results ('k') or by the results of the RUP/ANG clustering ('r')

5) ROSE DIAGRAM

ROSETYPE (4.18.1): enter '**s**' for symmetric (bidirectional) and 'a' for asymmetric (unidirectional) one

ROSEDIRECTION (4.18.2): enter '**s**' or 'd' to show strike or dip direction distribution on the rose plot

ROSEBINNING (4.18.3): enter 'a', 'b', '**c**' or 'd' for 2.5, 4.0, 10.0 and 22.5 deg bin sizes on the rose plot

6) OTHER

IDEALMOVEMENT (4.19): enter 'y' if you want to display ideal movement direction of slickensides, otherwise use '**n**'

LABELING (4.20): enter 'y' to label all of your data on the stereonet, otherwise use '**n**'

LINEWIDTH: used on stereonets in points; enter 1 - 9 for 0.1 – 0.9, or 0 for 1.0. Default **6** is for 0.6 points

COLORING (2): you had the chance to group everything by a user defined group code, by k-means clustering result and later by RUP/ANG clustering, and you had the chance to use color codes for different records, too. While coloring the output stereonets, either any codes are ignored ('**i**') to generate uniform black colors, or data records can be colored using their color codes ('c'). It is possible to color them using their user defined group code ('g'), using their the k-means clustering results ('k') or using their the results of the RUP/ANG clustering ('r')

GRAYSCALE (4.21): RGB colored ('**n**') and grayscale ('y') outputs are possible for the user's choice; grayscale output is recommended for publications

INPUT FILE FORMAT

XY DATA FORMAT

- XY data file (*.xy) is a tab separated text file with the following properties:
  - Has **4 columns**;
  - **First line is the header**, with arbitrary content (it is recommended to indicate column names here)

- 1st column must be **data point name**. The software will use this field to search for coordinates to merge with the original input (*.rgf) file content. In the case of repetition of data point names, error message will displayed

- 2nd and 3rd columns must be **coordinates** in user defined **numeric format** and order. Just correct numeric format will checked but otherwise not processed by the software just merged with the opened *.rgf file content. These values will overwrite the relevant rows of LOC_X and LOC_Y columns

- 4th column is for the (optional) **formation name**; the software will use this field to complete the original input (*.rgf) file content and overwrite the **FORMATION** field if it ha any record

	A	B	C	D
1	LOCATION	LOC_X	LOC_Y	FORMATION
2	001	35.167340	44.696108	FARS
3	002	35.060867	44.589635	FARS
4	003	34.689926	44.218694	FARS
5	004	35.323870	44.852638	FARS
6	005	34.704831	44.233599	FARS
7	006	34.589101	44.117869	FARS
8	007	34.833382	44.362150	FARS
9	008	34.840982	44.369750	FARS
10	009	34.724930	44.253698	FARS
11	230	35.168929	44.697697	BEKHME
12	ANG001	35.169656	44.698424	BEKHME
13	ANG002	35.171196	44.699964	BAKHTIARI
14	ANG003	35.173050	44.701818	KAREEM
15	ANG004	36.164102	45.692870	KAREEM
16	ANG005	35.769995	45.298763	KAREEM
17	ANG006	35.758512	45.287280	KAREEM
18	ANG007	35.708102	45.236870	KAREEM
19	ANG008	35.708104	45.236872	KAREEM
20	ANG009	36.457510	45.986278	KAREEM
21	ANG010	35.708932	45.237700	KAREEM
22	ANG011	35.725867	45.254635	KAREEM
23	BE-03	35.742802	45.271570	QAMCHUQA
24	BE-04	35.759736	45.288504	QAMCHUQA
25	BE-05	35.776671	45.305439	QAMCHUQA
26	BE-06	35.793606	45.322374	QAMCHUQA
27	BE-07	35.810541	45.339309	BEKHME
28	BE-08	35.827475	45.356243	BEKHME
29	BE-09	35.844410	45.373178	BEKHME

- TRJ data file (*.trj) is a tab separated text file with the following properties:
  - Has **3 columns**
  - **First line is the header**, with arbitrary content (it is recommended to indicate column names here)

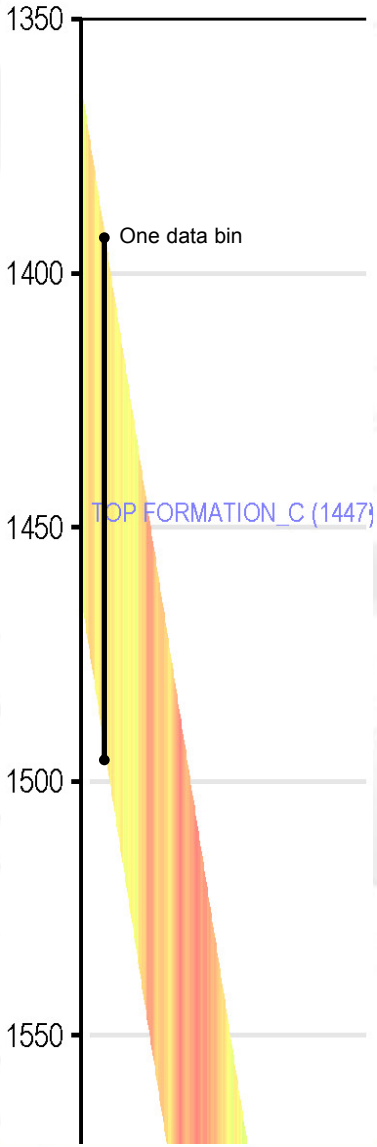
- 1st column is for the **data point depth**. Only numeric values are accepted here

- 2nd and 3rd columns are for **coordinates** in user defined **numeric format** and order. The following data conventions are accepted here:
  - azimuth – plunge data set: the 2nd row has numeric values less than 360 but greater than 0, and the 3rd row has numeric values less than 90 but greater than 0
  - X – Y: any numbers in the coordinate data format of the project

The software identifies **automatically** which data convention is in use. If the content of the trajectory file fulfills the first criteria, then it will be handled as an azimuth – plunge data set, otherwise it will be handled as an X – Y coordinate pair

	A	B	C
1	DEPTH	LOC_X	LOC_Y
2	1181.27	35.167340	44.696108
3	1198.57	35.060867	44.589635
4	1210.65	34.689926	44.218694
5	1212.55	35.323870	44.852638
6	1215.19	34.704831	44.233599
7	1217.23	34.589101	44.117869
8	1217.55	34.833382	44.362150
9	1221.93	34.840982	44.369750
10	1229.11	34.724930	44.253698
11	1231.16	35.168929	44.697697
12	1233.14	35.169656	44.698424
13	1241.48	35.171196	44.699964
14	1242.06	35.173050	44.701818
15	1243.51	36.164102	45.692870
16	1252.33	35.769995	45.298763
17	1253.15	35.758512	45.287280
18	1265.42	35.708102	45.236870
19	1265.97	35.708104	45.236872
20	1280.33	36.457510	45.986278
21	1281.85	35.708932	45.237700
22	1292.46	35.725867	45.254635
23	1300.51	35.742802	45.271570
24	1303.99	35.759736	45.288504
25	1312.55	35.776671	45.305439
26	1314.06	35.793606	45.322374
27	1330.94	35.810541	45.339309
28	1335.92	35.827475	45.356243
29	1341.43	35.844410	45.373178

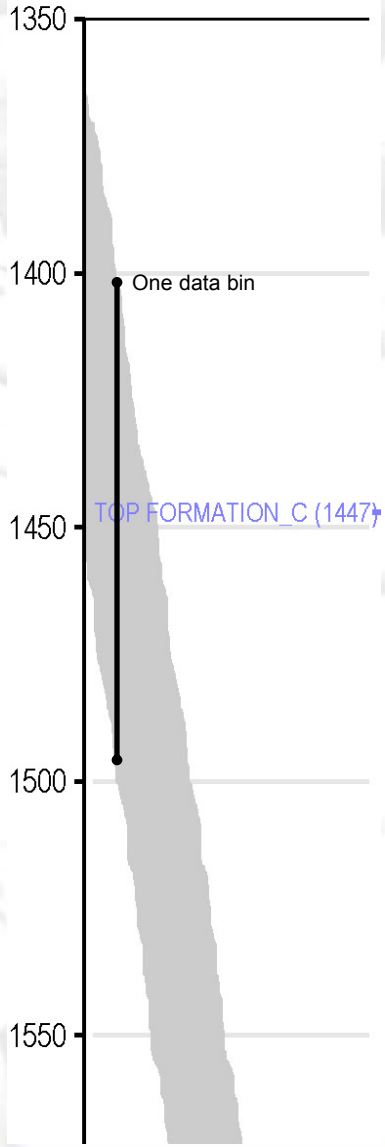
PROCESSING AND DISPLAY



- While well data processing, a **moving average curve** is calculated to highlight changes in the features geometry and/or frequency
- In order to do a moving average calculation, a processing interval (**=bin**) and the step have to be defined
- Bin size can be given by the user either in **meters** (e.g. 50m) or in **pieces** (e.g. 100 data). Different settings are resulting in different data bins, which generates different results
- Step size is default: 1m if the bin size is in meters, and 1 piece of data if the interval length is defined by data number

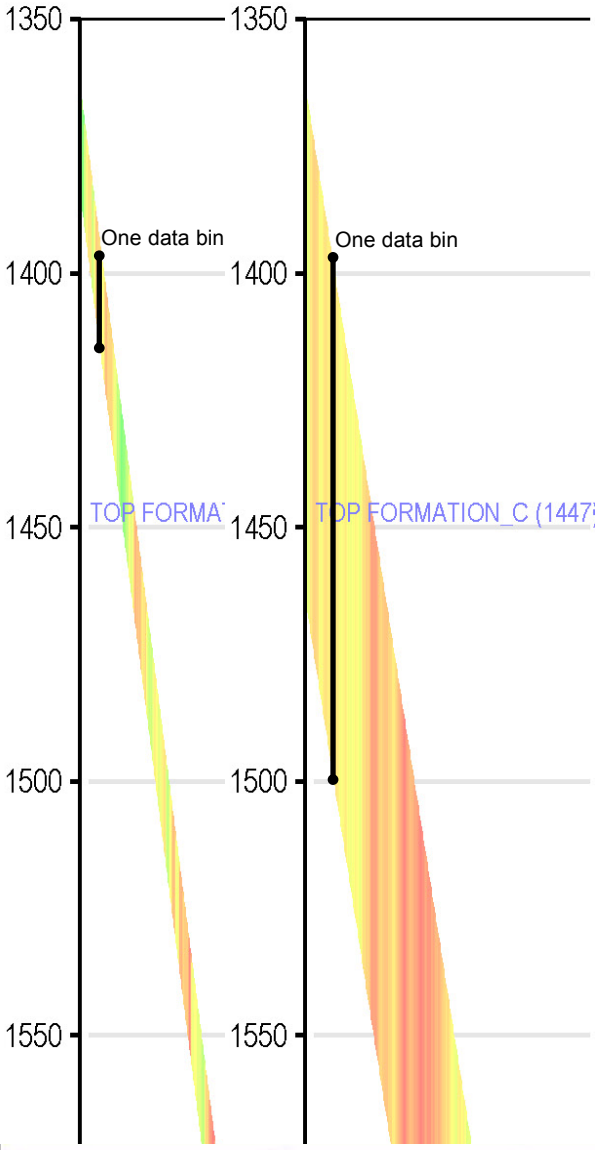
⬅ Test data set, with **100m data bin value**. Step interval during moving averages calculation is 1m. Vertical colored bars are showing each bin intervals . Color of the bars refers to the relative data number (green: many data, red: less data)

Test data set, with **100 piece data bin value**. Step interval during moving averages calculation is 1 piece of data. Vertical gray bars are showing each bin intervals. No coloring: each bins have identical number of data ➡



PROCESSING AND DISPLAY

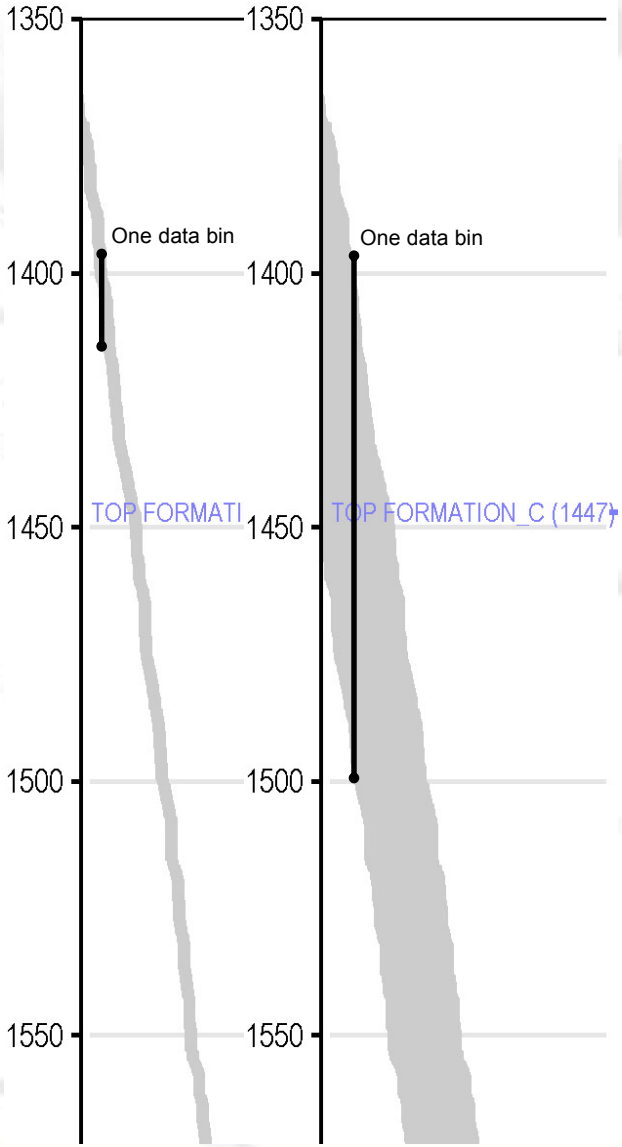
WELL INTERVAL LENGTH I.



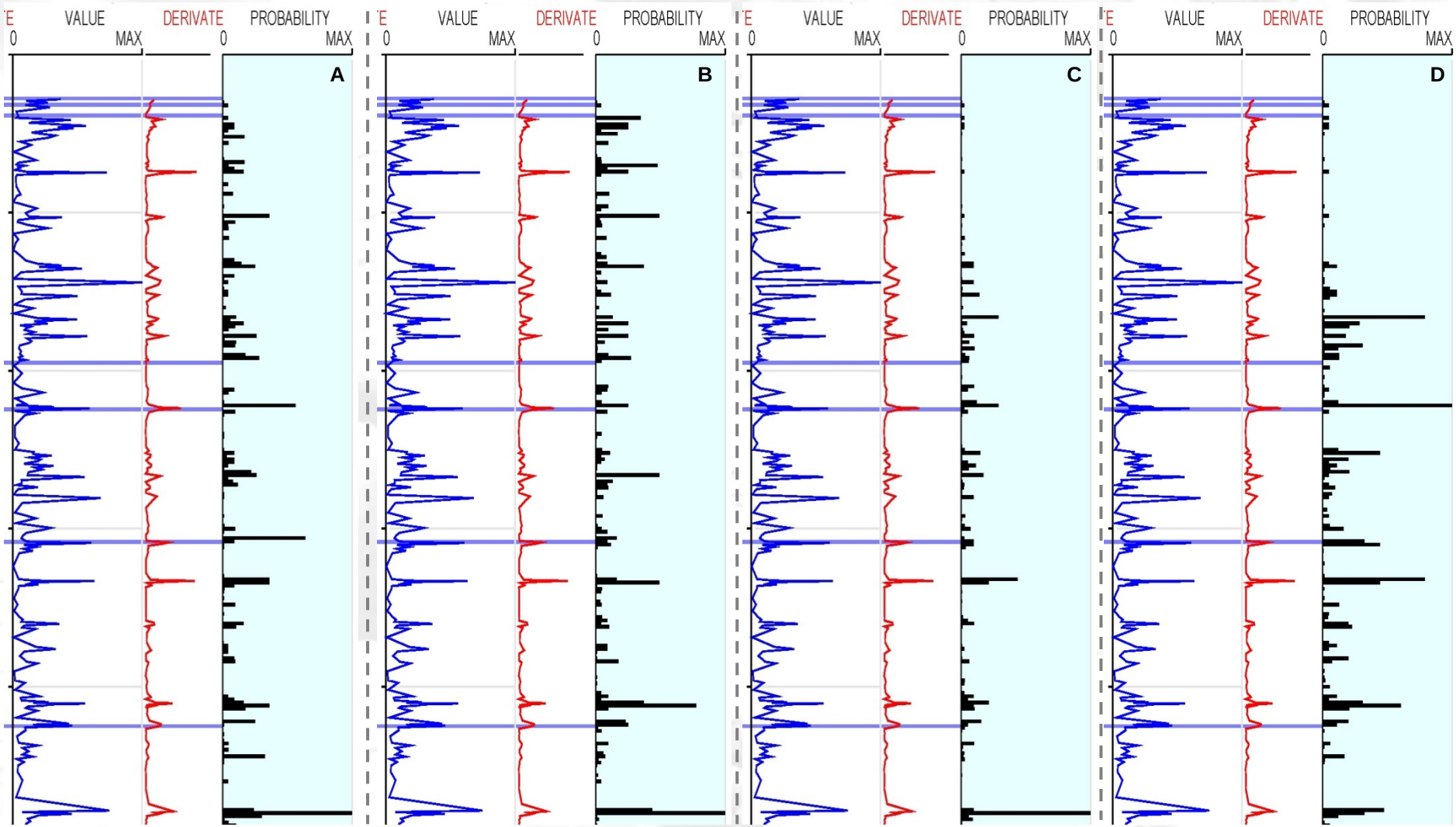
- During the moving average calculation, the bin size can be changed by the user
- Bin size can change on a 0 - 5000m or 0 - 5000 piece of data scale
- Using too small bins, maybe you will see a noisy outcome because 'over zooming' the data set
- Using too large bins, you can easily hide the real changes because having less resolution than required
- It is proposed to try to do many processing loops, varying both the bin size and bin type (meters or piece of data) to study the same data set with different resolutions

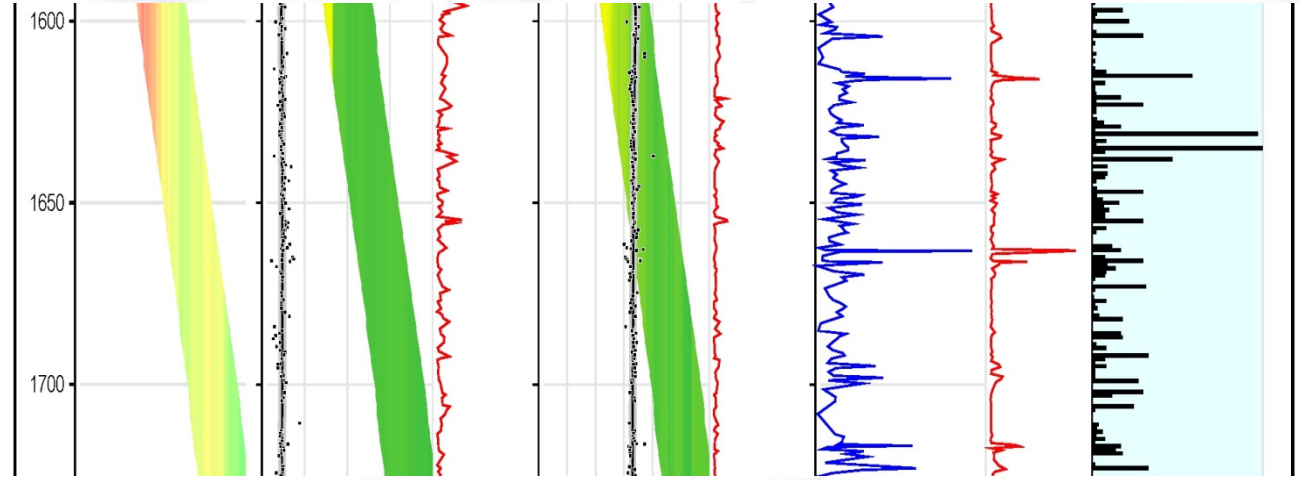
⬅ Test data set, with 20m and 100m data bin value. Step interval during moving averages calculation is 1m. Vertical colored bars are showing each bin intervals. Please compare the intervals taken into account while processing

Test data set, with 20 and 100 piece data bin value. Step interval during moving averages calculation is 1 piece of data. Vertical gray bars are showing each bin intervals. Please compare the intervals taken into account while processing ➡

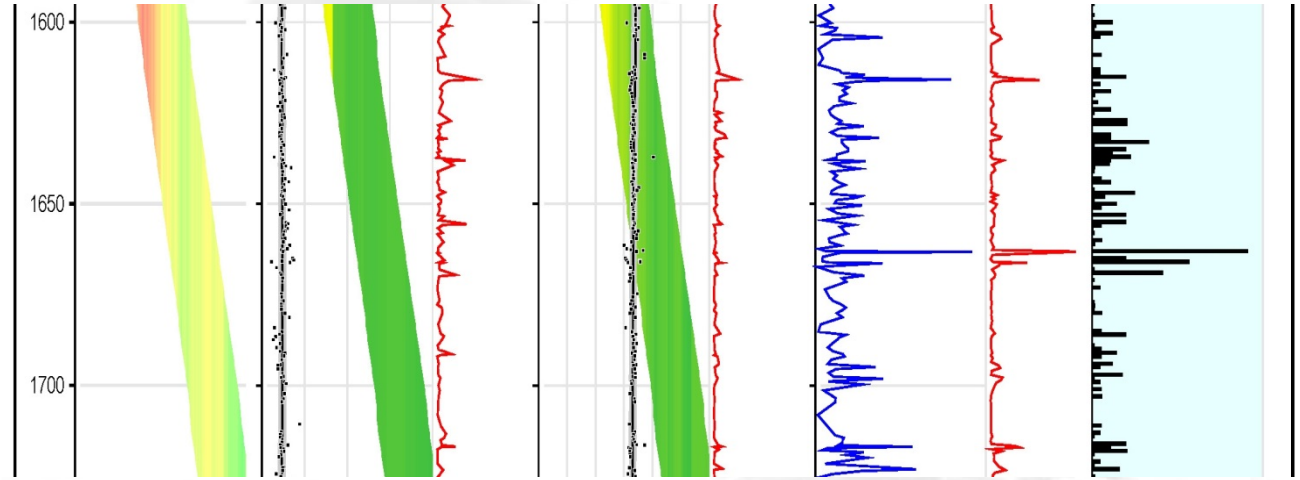


Fault / deformation zone reliability graph for the same data set, using 20m and 100m, and 2b0 and 100 piece data bin size, respectively
Please note the same data set was processed using different settings





Test data set, with 100 piece of data bin value. Bin interval middle point is the **average depth** of the binned data



Test data set, with 100 piece of data bin value. Bin interval middle point is the **median depth** of the binned data

- During the moving average calculation, one of the questions is about the **middle point depth** of the used bins
- In the case of a **dense** and **homogeneous** set, the **arithmetic average** depth of the binned data can be used
- In the case of **inhomogeneous** data set, and especially having significant **gaps**, **median depth** of the binned data is recommended to use

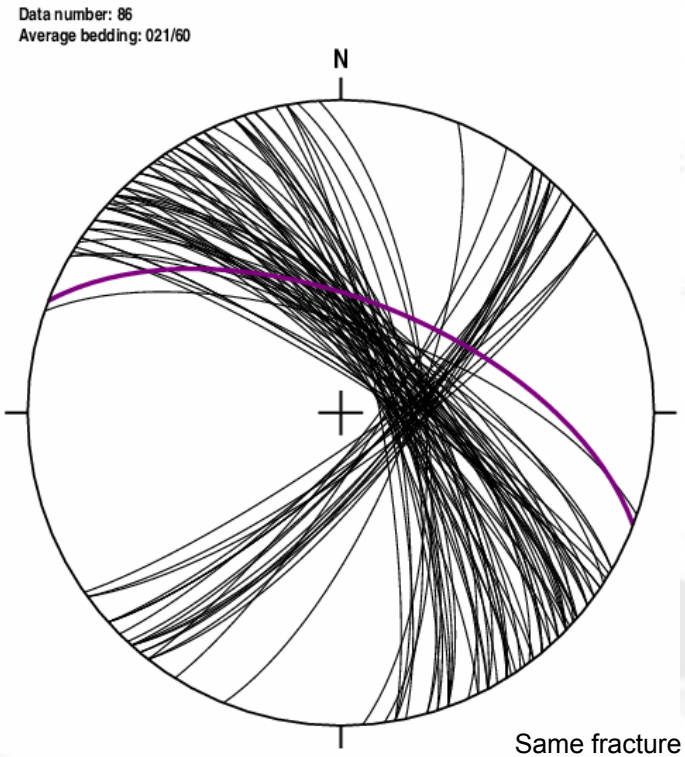
GERMAN DIP DIRECTION AND RIGHT HAND RULE

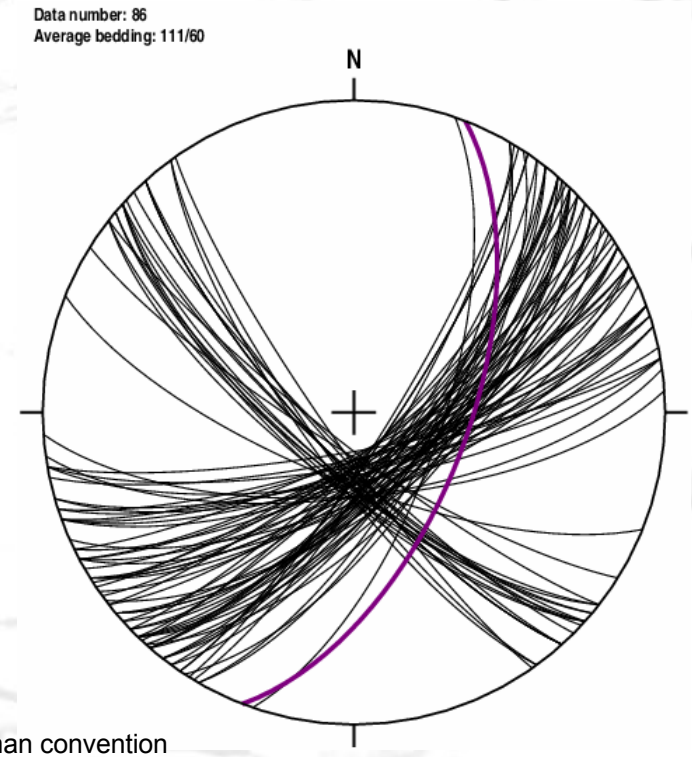
DIP DIRECTION

- In this case the planes and lineations are described by the **dip direction** and **dip angle** of the element, in this order
- To convert German convention data set to right hand rule data, add -90 degrees to the dip direction

RIGHT HAND RULE

- In the case of right hand rule, planes are described with **strike** direction (-90 degrees from dip direction) and **dip angle** of the planar object, in this order
- Lineations are presented by **bearing and plunge**, which are the **equivalents** of dip direction and dip angle





Same fracture data set processed as German convention data (left) and right hand rule data (right)

PROCESSING AND DISPLAY

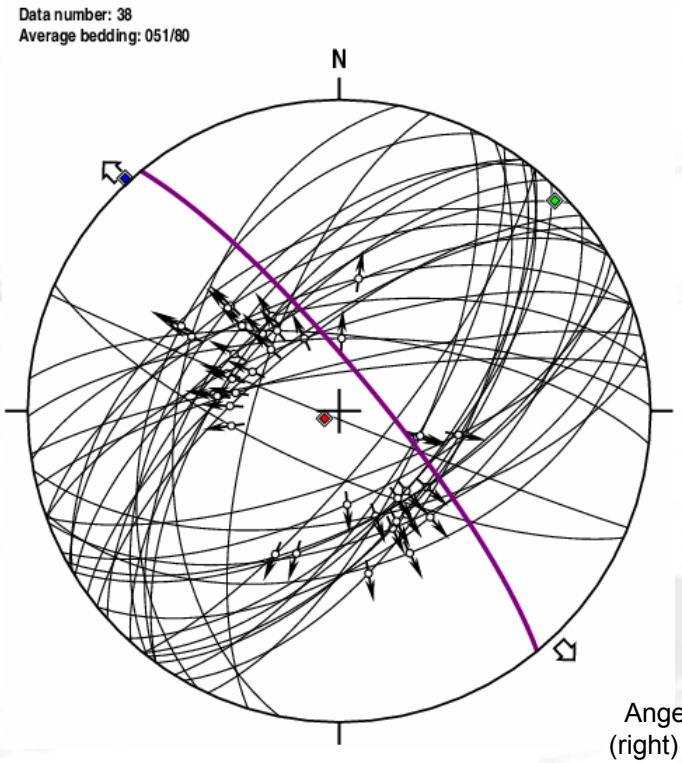
ANGELIER AND HOEPPNER PLOT

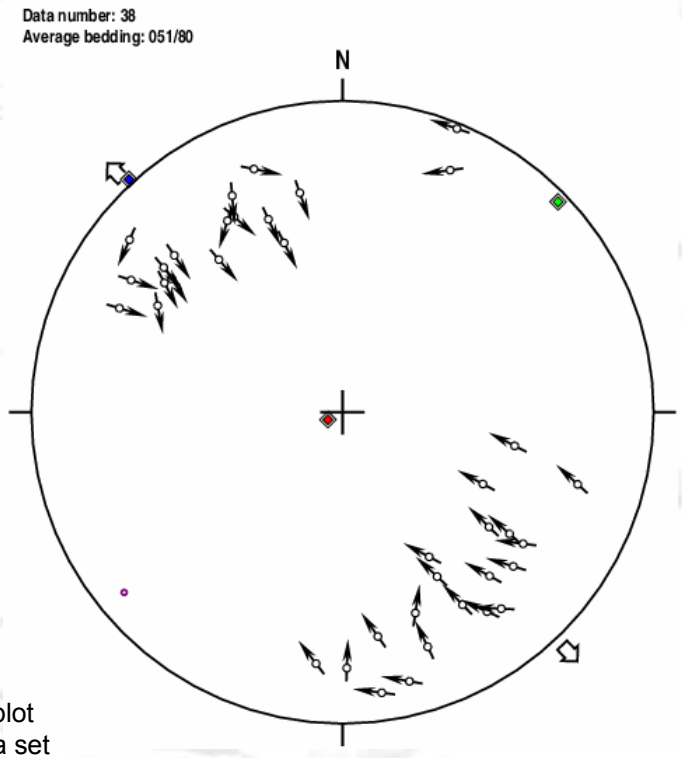
ANGELIER PLOT

- Angelier plot shows all planes as **great circles**, and all lineations as **points**
- In the case of slickenside visualization, fault plane and the **movement direction** along it is displayed
- Easy to read in the case of **homogenous** data sets, or in the case of **small** number of data

HOEPPNER PLOT

- Hoeppner plot shows **pole point of the planes**
- Lineations are displayed as **points** as well
- Striae are plotted as **pole points** pointing in the **movement direction of the hanging wall**
- This representation is really useful in the case of **large** or **inhomogeneous** data sets





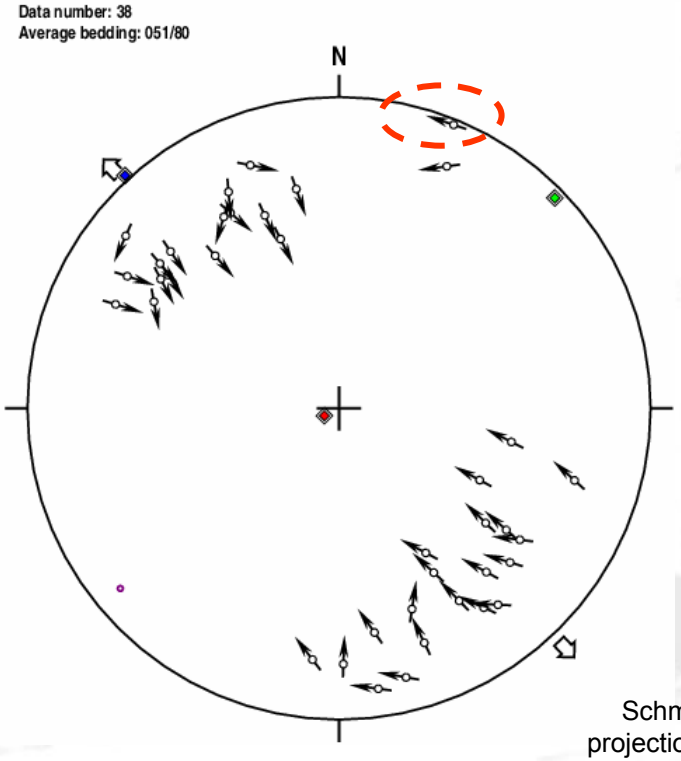
Angelier plot (left) and Hoeppner plot (right) of Angelier's (1990) TYM data set

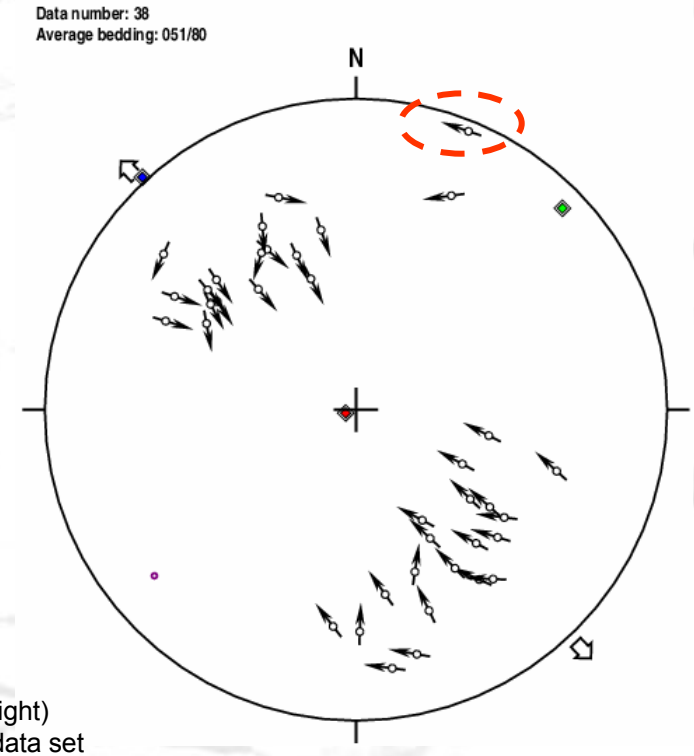
SCHMIDT NET

- Equal area projection
- Great circles are **Bezier** curves

WULFF NET

- Equal angle projection
- Great circles are **arcs**





Schmidt-net (left) and Wulff-net (right)
projection of Angelier's (1990) AVB data set

PROCESSING AND DISPLAY

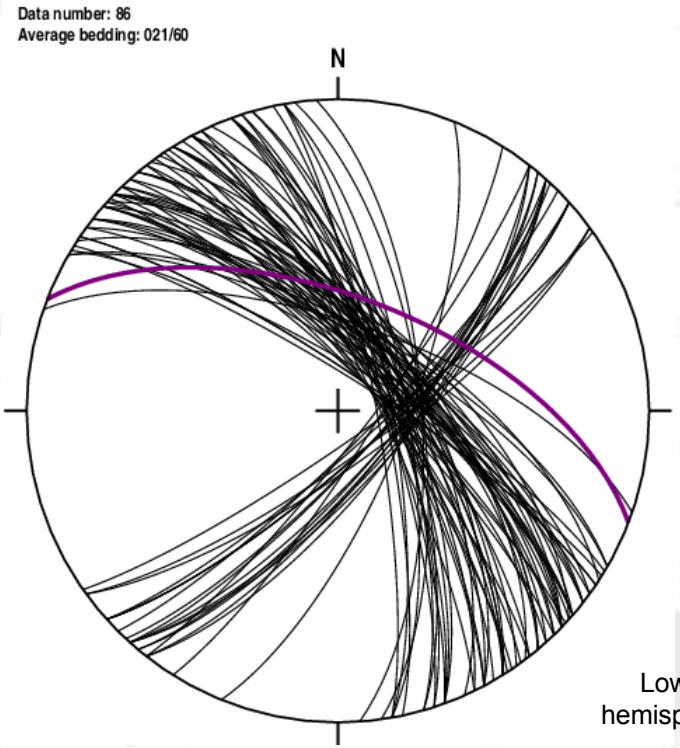
LOWER AND UPPER HEMISPHERE

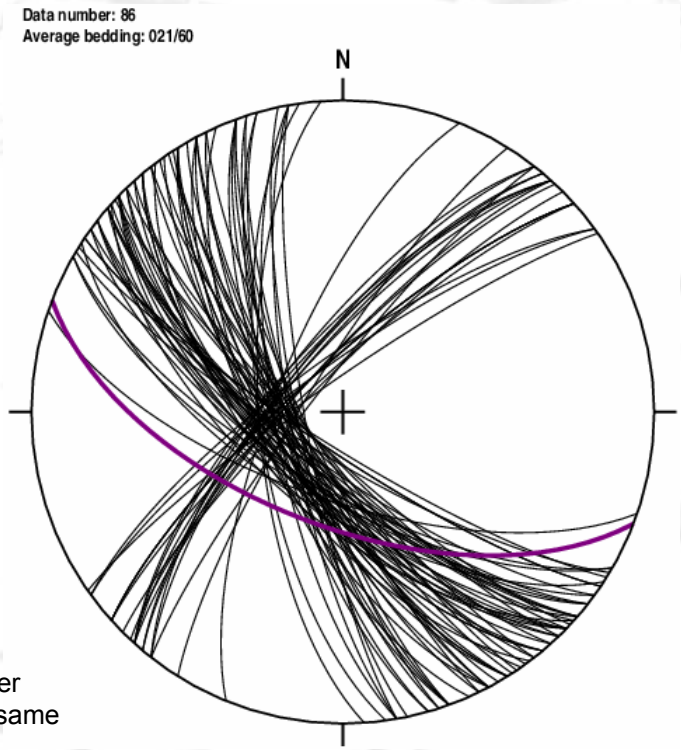
LOWER HEMISPHERE

- Projects on **lower** hemisphere

UPPER HEMISPHERE

- Projects on **upper** hemisphere

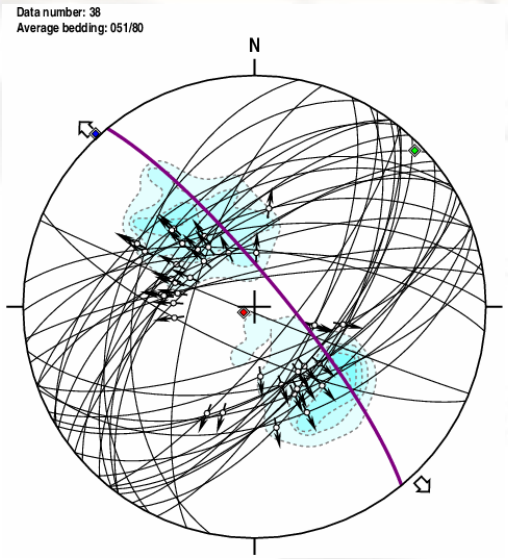




Lower hemisphere (left) and upper hemisphere (right) projection of the same inverse slickenside data set

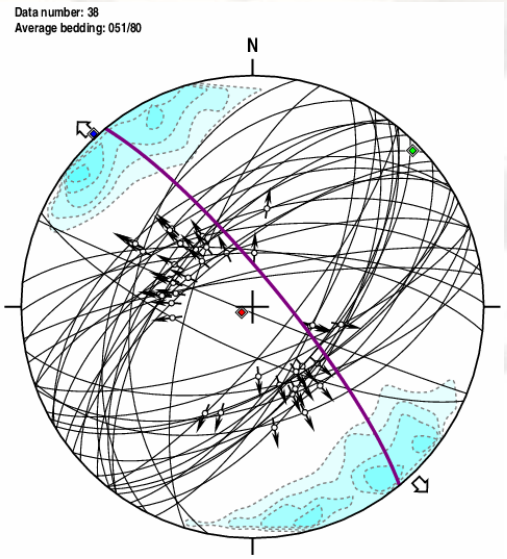
PROCESSING AND DISPLAY

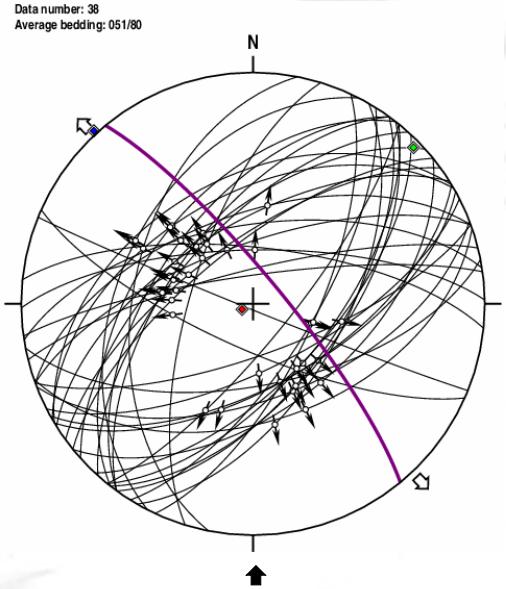
CONTOURING



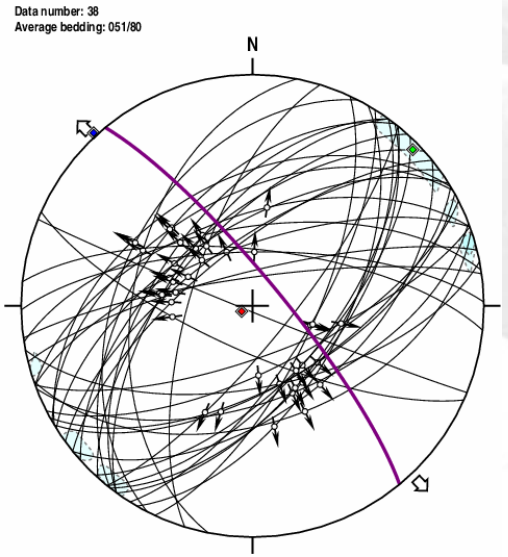
Contouring of the input data set using plane **dip direction** or lineation **bearing** as input

Contouring of the input data set using **plane normals** or **lineation bearing** as input



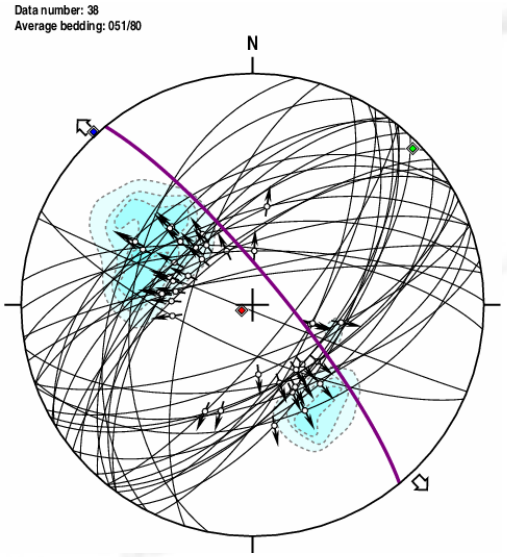


Original input data set, **no contouring** method was chosen



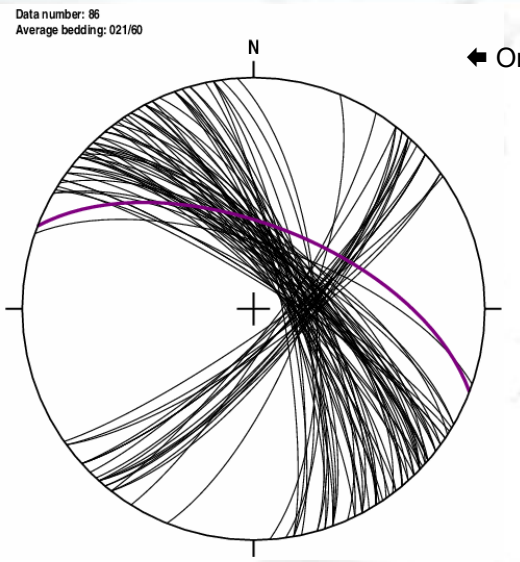
Contouring of the input data set using plane **strike** direction or lineation **bearing** as input

Contouring of the input data set using **striae bearing** or **lineation bearing** as input



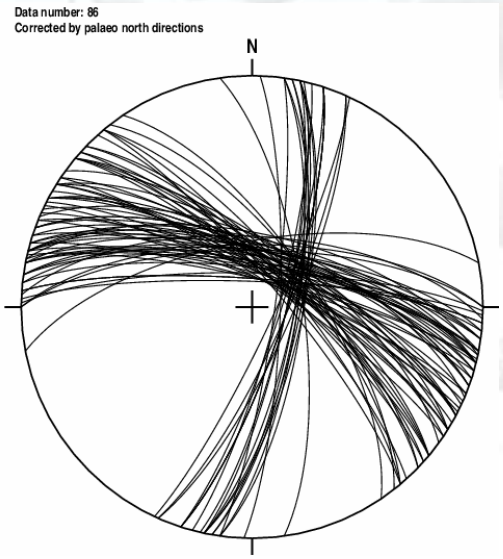
PROCESSING AND DISPLAY

CORRECTION WITH AVERAGE BEDDING AND PALAEONORTH DIRECTION

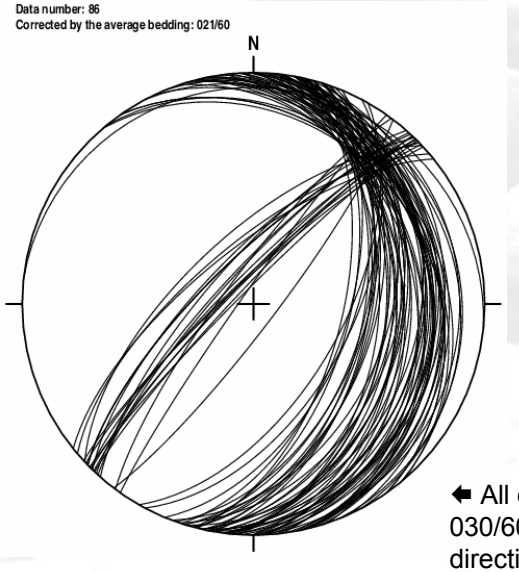


⬅ Original dataset

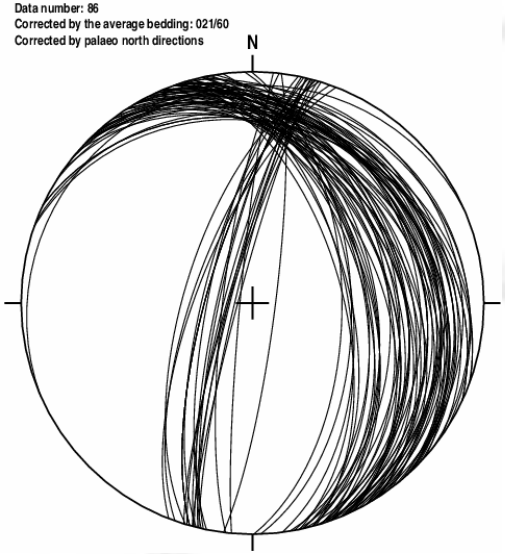
Corrected by 30 deg palaeonorth direction for all of data ➡



- Selecting **bedding correction** option, all data will be re-tilted by the **average bedding** measured in the outcrop. If no bedding data, the plot for measured and corrected data will be the **same**
- Selecting **palaeonorth direction correction**, all data will be corrected with the palaeonorth direction entered in RGF file. Correction will be done **record by record**
- Third possibility is to apply **both** of these correction re-tilting everything by average bedding plane direction **and** by palaeonorth direction



⬅ All data corrected by 030/60 average bedding direction



Corrected by 030/60 average bedding and by 30 deg palaeonorth direction ➡

INVERSION I.

REGRESSION WITH A PRIORI *ASSUMPTION*

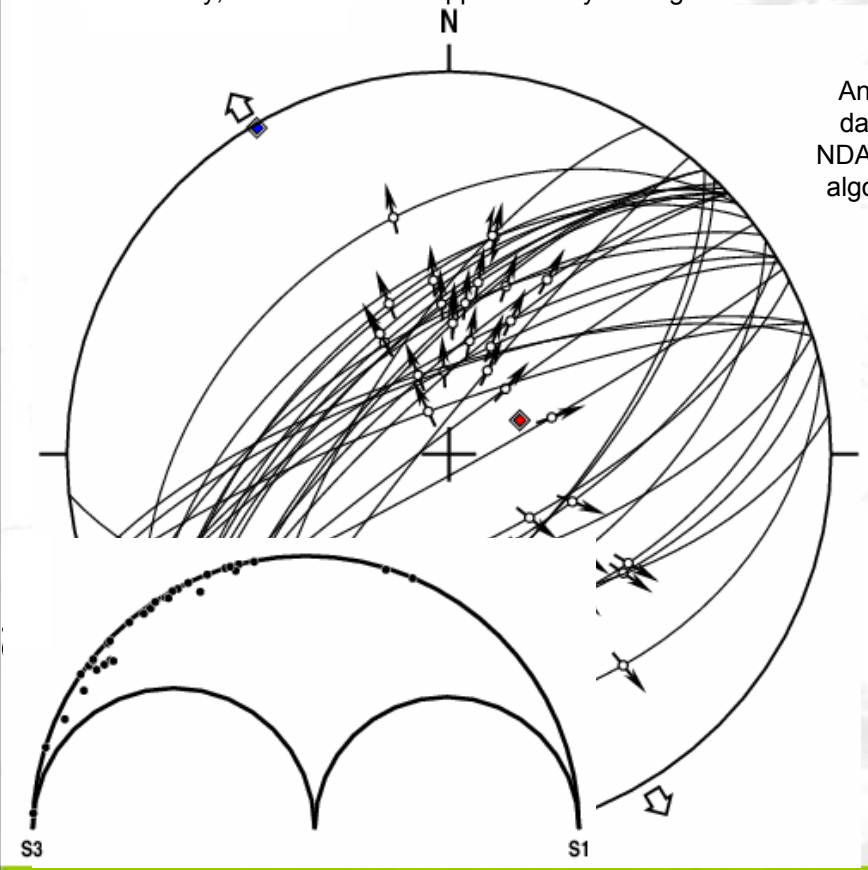
- Regression methodologies with a priori assumption are calculating maximum and minimum stress orientations for each slickenside **based on fault plane and striae geometry**, and searching for the best fitting stress tensor
- To calculate stress axes direction, **a priori assumption is required** about the **a angle** between the fault plane (holding the slickenside) and the maximum stress axis direction
- Usually, a value could be approximately 30 degrees

NDA after Spang (1972)

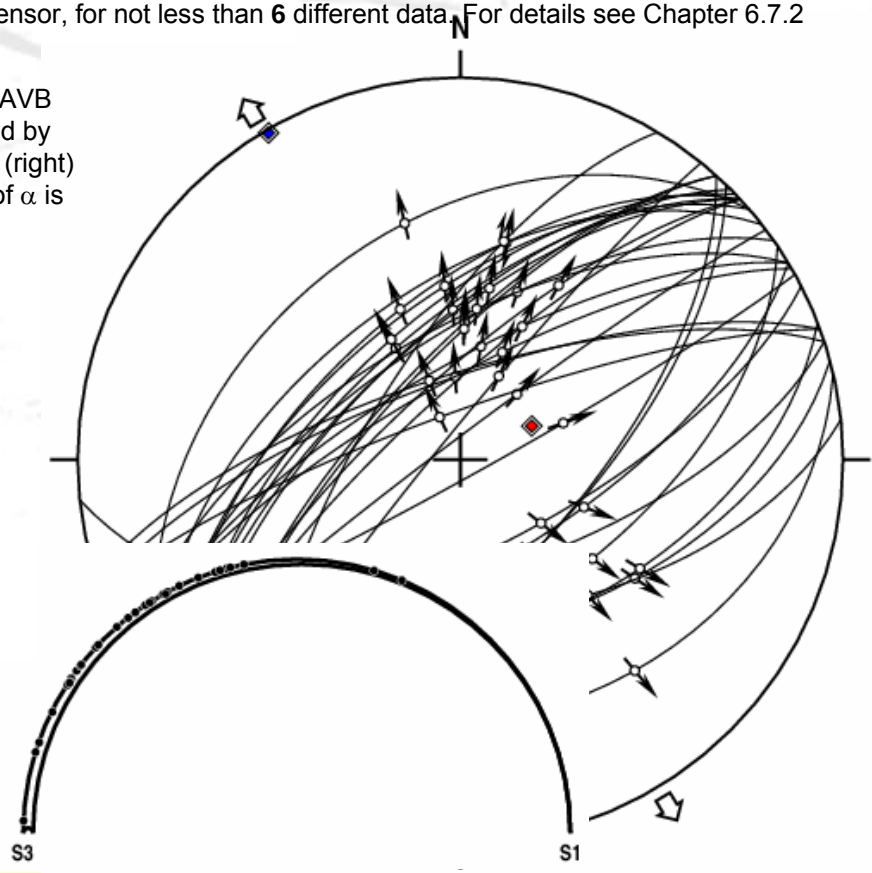
- NDA methodology calculates the best fitting stress tensor on this way for the **entire** data set with not less than **6** different data. For details see Chapter 6.7.1

PTN after Turner (1953)

- PTN methodology calculates the best fitting tensor for **maximum, minimum and intermediate stress orientations** and processes this tensor, for not less than **6** different data. For details see Chapter 6.7.2



Angelier's (1990) AVB data set processed by NDA (left) and PTN (right) algorithms. Value of α is 30 degrees



INVERSION II.

REGRESSION WITHOUT A PRIORI *ASSUMPTIONS*

- Inversion methodologies without a priori assumptions satisfying Angelier's (1979) basic equation about **no traction in direction perpendicular to striae direction**; b_i is perpendicular to the striation, T is the stress tensor and n_i is the plane normal of the i^{th} data

$$b_i \cdot T \cdot n_i = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = 0$$

FRY (1999)

- Fry's (1990) methodology is a classic six dimensional regression to solve the equation on the left hand side; use for minimum **6** data at the same set (see Chapter 6.7.3)

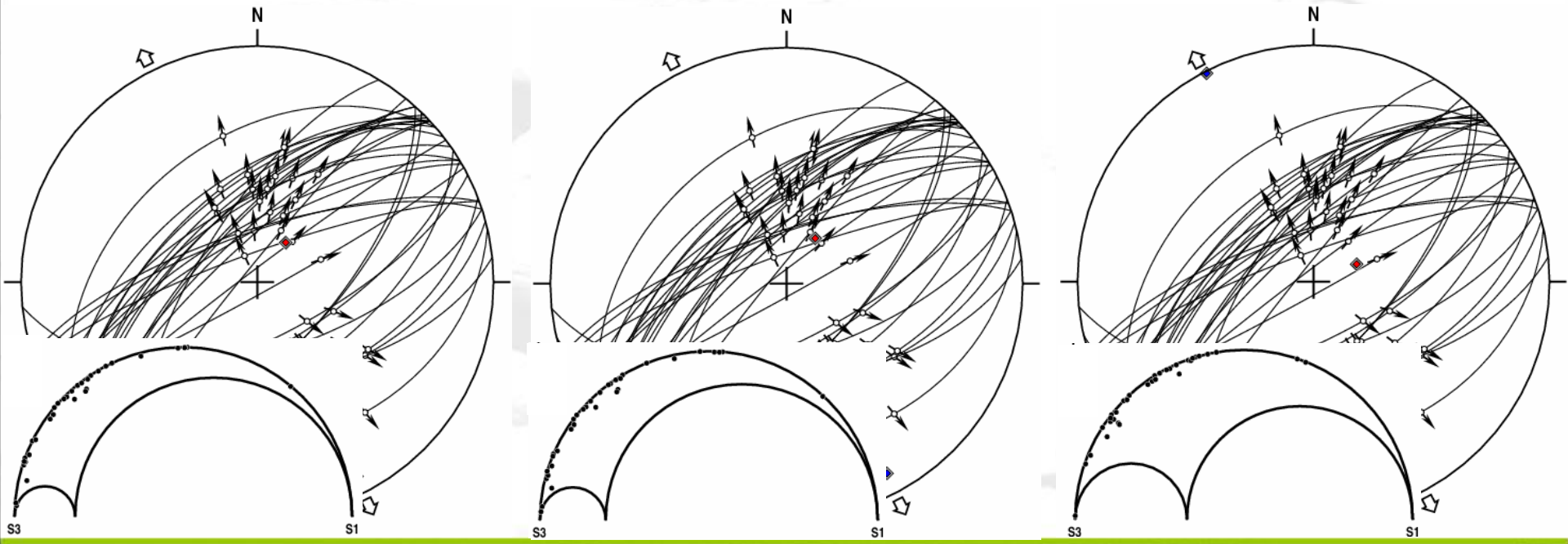
SHAN ET AL (2003)

- Shan et al's (2003) solution is five dimensional regression assuming $s_{11} + s_{22} + s_{33} = 0$; use for minimum **5** data at the same set (see Chapter 6.7.4)

MICHAEL (1984)

- Michael's (1984) solution is to calculate the shear vector with assumed unit length (see Chapter 6.7.5) and find the best fitting stress tensor using five dimensional regression for no less than **6** different data

Angelier's (1990) AVB data set processed by Fry's (1999), Shan et al's (2003) and Michael's (1984) regression methodologies, respectively



INVERSION III.

DIRECT INVERSION, ITERATIVE DIRECT INVERSION

- Direct inversion technologies are expressing an **estimator** describing the cumulative misfit between the measured and computed slickenside movement direction (misfit angle, divergence vector, etc)
- Basic assumption after Angelier (1979) is to suppose $s_{11} + s_{22} + s_{33} = 0$ and $s_{11}^2 + s_{22}^2 + s_{33}^2 = 3/2$; the stress tensor below fits to these criteria;
- To minimize the misfit, create **partial derivates** of the stress estimator, and solve the equations

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} \cos(\psi) & \alpha & \gamma \\ \alpha & \cos\left(\psi + \frac{2\pi}{3}\right) & \beta \\ \gamma & \beta & \cos\left(\psi + \frac{4\pi}{3}\right) \end{bmatrix}$$

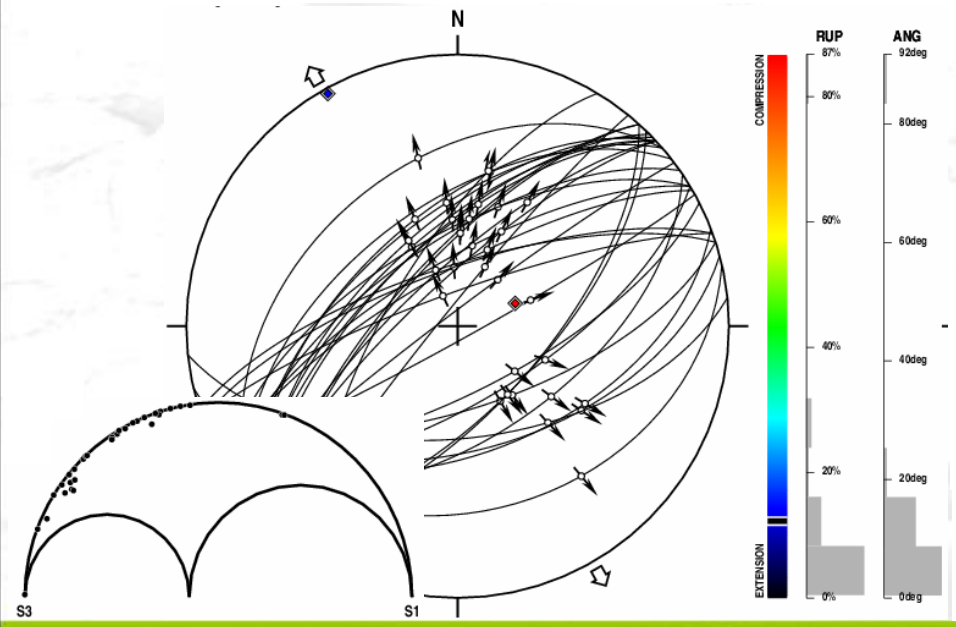
ANGELIER (1990)

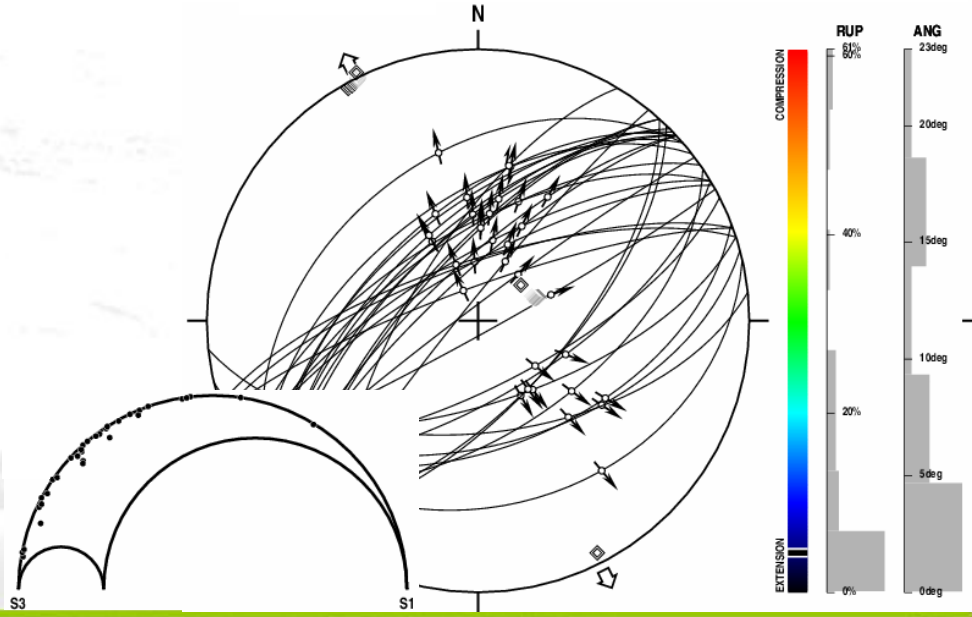
- Angelier's (1990) methodology calculates a **misfit vector** ν between the **measured** and **calculated** shear vector, and minimizes its length (see Chapter 6.7.6) for not less than **4** different data

MOSTAFA (2005)

- Because the measured stress vector length is **unknown**, Angelier (1990) assumed constant measured shear vector length λ; Mostafa's (2005) methodology is the same as Angelier's (1990) one, with **iteration in the measured shear vector length λ** to decrease ν misfit vector length (see relative upsilon RUP bars below, and Chapter 6.7.7) for not less than 4 different data

⬇ Angelier's (1990) AVB data set processed by Angelier's (1990) and Mostafa's (2005) methodology ⬇

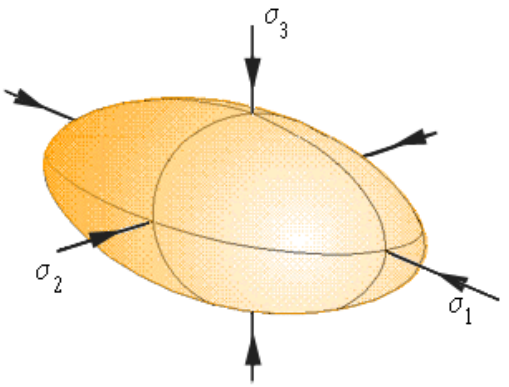


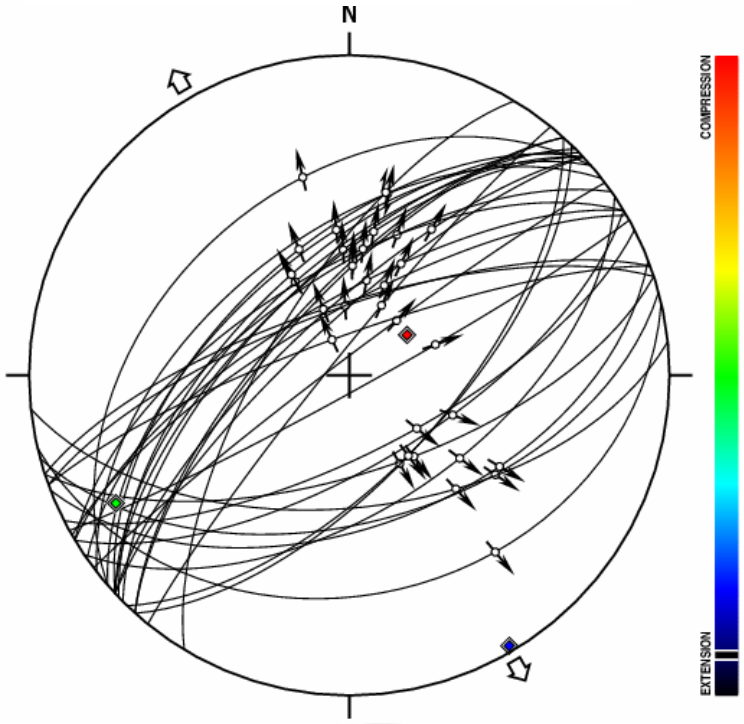


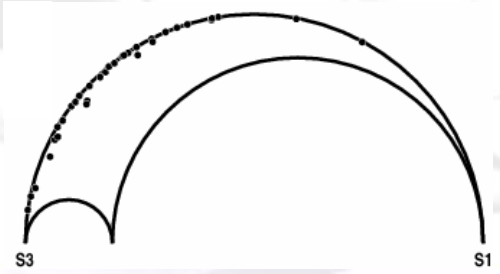
PROCESSING AND DISPLAY

INVERSION IV. – BRUTE FORCE INVERSION

- Brute force methodologies are really simple but time consuming technologies to find the best fitting stress tensor of the investigated data set with not less than **6** different data
- Brute force inversion is a **simple iteration** with **no preliminary assumption** and with **no heuristic approach**
- This methodology has a 'trial and error' philosophy – the stress tensor with the **smallest cumulative misfit angle** (ANG) is the 'best fitting' stress tensor
- Because the stress ellipsoid is characterised by the **stress axes direction** (perpendicular to each other) and the **stress axes length** (or stress ellipsoid shape), a number of stress tensor candidates with different geometry and orientation can be used
- **Iterating** in these variables, and '**zooming in**' on the local maxima, it is possible to find the best fitting stress tensor – this will be really the best fitting one, because all theoretical solutions were cross checked







PROCESSING AND DISPLAY

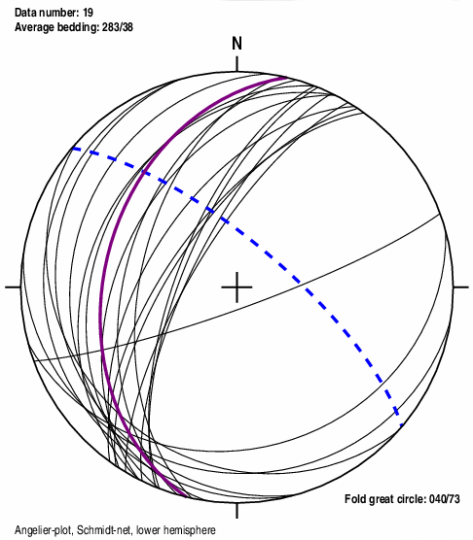
BINGHAM STATISTICS: FOLD GREAT CIRCLE, TRANSPORT DIRECTION

BINGHAM STATISTICS

- Average computation can be carried out on different ways; one of them is to calculate **arithmetic average**, or median
- In the case of **directional data** (e.g. stress vector direction) this methodology might be misleading (average of compression vectors 040/10 and 220/10 is ~220/90 and not 040/00 or 220/00)
- One of the possible solution is the **regression** finding the best fitting 'average' directions of the processed data set

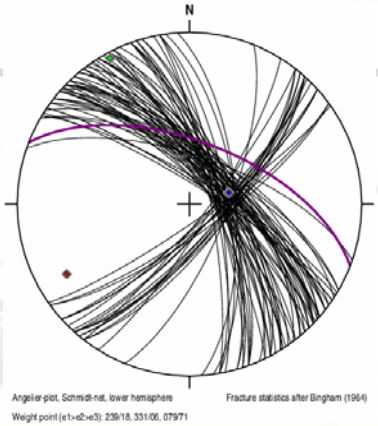
FOLD GREAT CIRCLE CALCULATION

- This methodology is used to find the best fitting plane of 'foldsurface' plane normals to calculate best fitting planes: the great fold circle



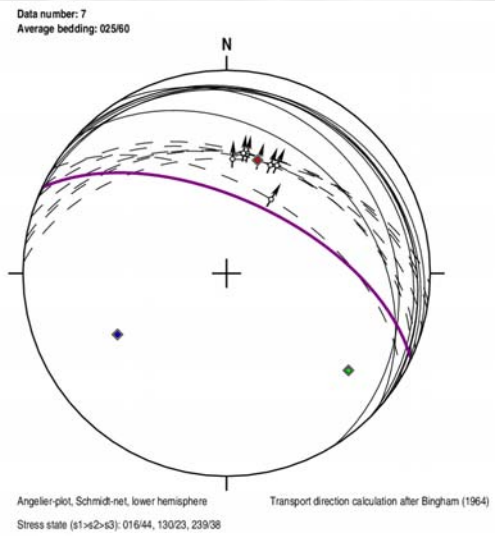
Fold great circle of a 'foldsurface' data set using Bingham statistics to find the best fitting plane

- Bingham directional statistics calculates **directional density** of the input fracture data set, resulting the maximum, intermediate and minimum density directions



MEAN TRANSPORT DIRECTION CALCULATION

- To calculate the best fitting transport direction for SC data, this methodology is used instead of the classic paleo stress inversion

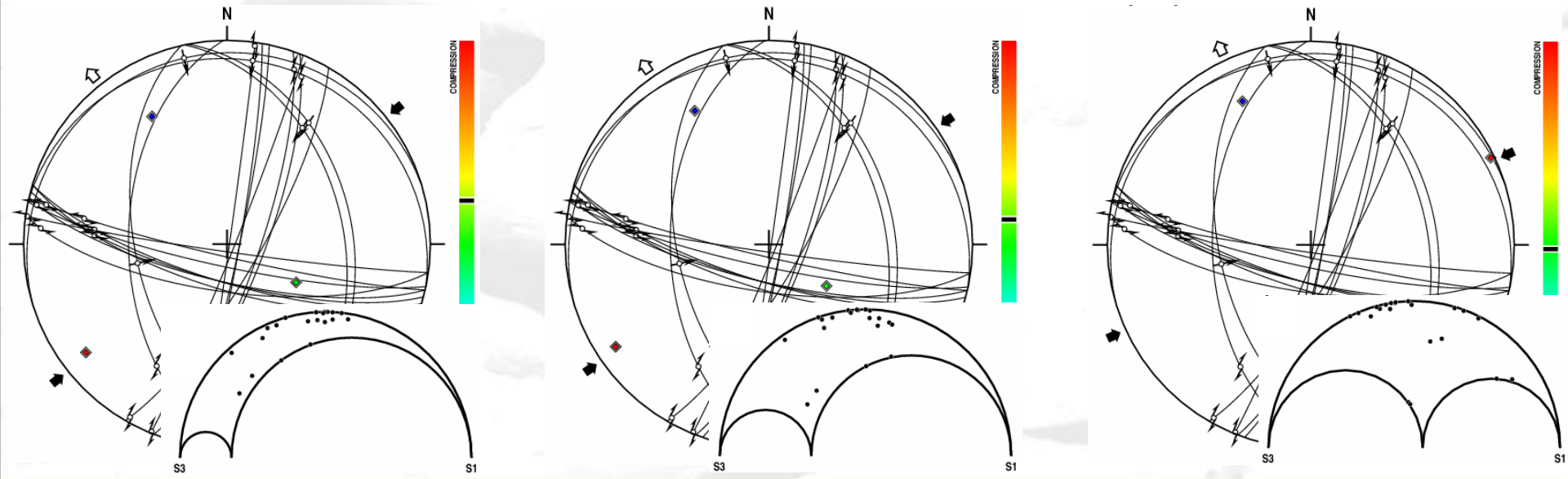


Result of mean transport direction calculation for SC data

Bingham statistics result for a fracture set

- Regression technologies with **a priori** assumptions are using **a angle** to describe angle between the movement direction along the fault plane and the maximum stress direction
- This value can be constant, or the best fitting value can calculate using iteration methodologies (see Ortner et al. 2002)

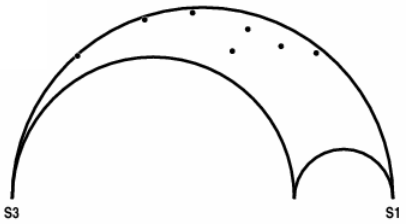
Same data set processed with NDA technology using different a angles between movement direction and maximum stress axis direction, 15, 30 and 70 degrees, respectively



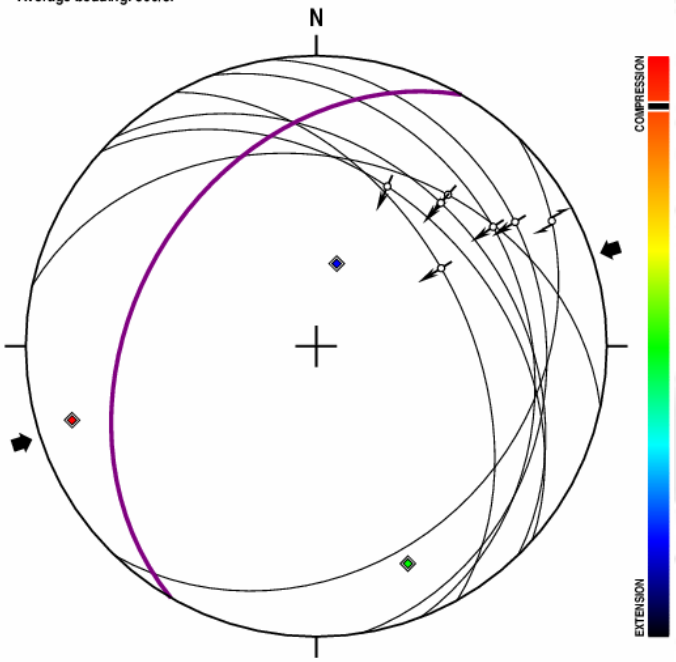
PROCESSING AND DISPLAY

FORCED ANDERSONIAN STRESS CONDITIONS

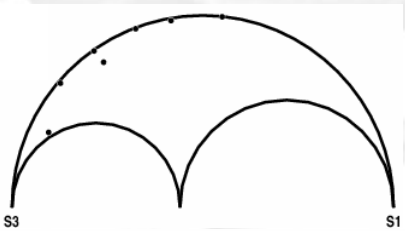
- Some inversion methodologies are sensible on the **symmetry** of the input data set; in some case of asymmetry the result is mathematically correct, but not from geological point of view
- One of the possible solutions could be to generate a **virtual, symmetrical data set**; this virtual set is composed by the original data rotated by **180 degrees** around a vertical axis
- The result stress field axes will be definitely **Andersonian** with two horizontal and one vertical stress axes



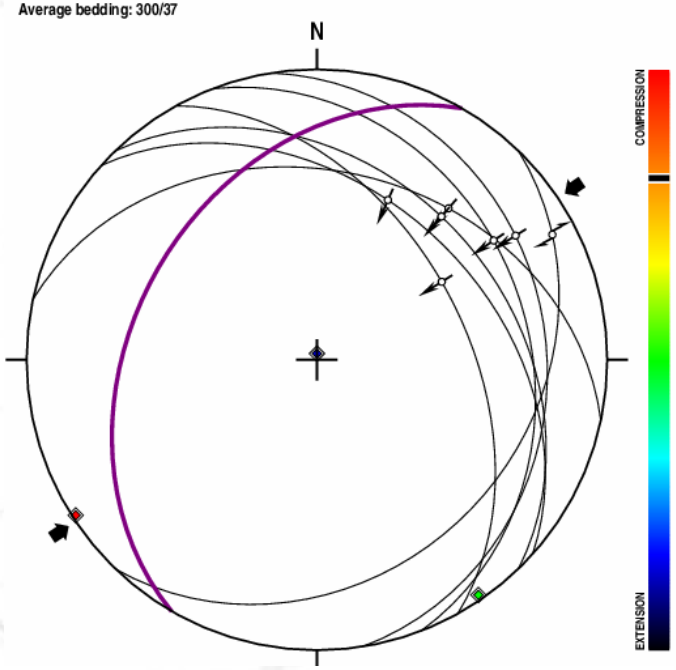
Data number: 7
Average bedding: 300/37



Same data set processed **without** (left) and **with** (right) forced Andersonian stress conditions, using virtual symmetric data set

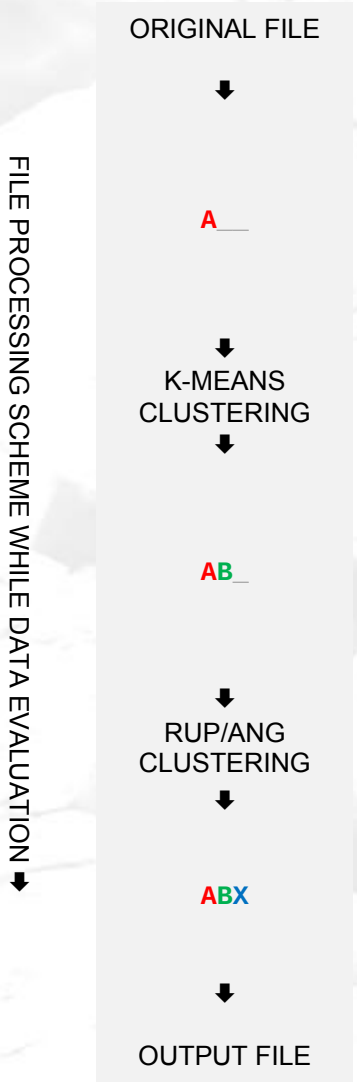


Data number: 7
Average bedding: 300/37



PROCESSING AND DISPLAY

IMPORTING GROUPS



- While opening, it is possible to choose which digit of the group code to import as a user defined group code

1. Importing the **first** character (user defined group), you will keep the original user defined groups of the file you have been created for a while

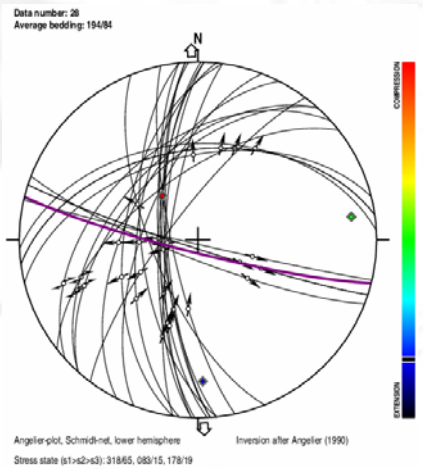
2. Importing the **second** character (k-means clustering result), you will keep the original k-means clustering result groups and use them as user defined group code

3. Importing the **third** character (RUP/ANG clustering result), you will keep the original RUP/ANG clustering result groups and use them as user defined group code

PROCESSING AND DISPLAY

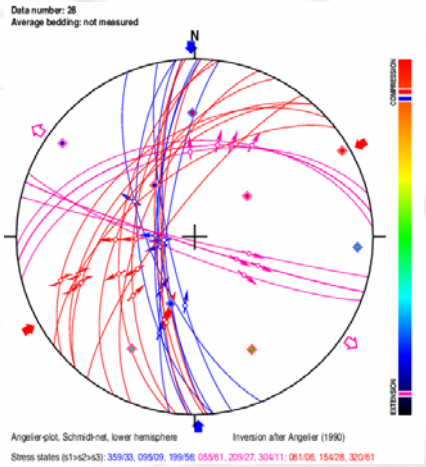
PROCESSING WITH AND WITHOUT USING GROUPS

USER DEFINED GROUPS **NOT USING** ➡



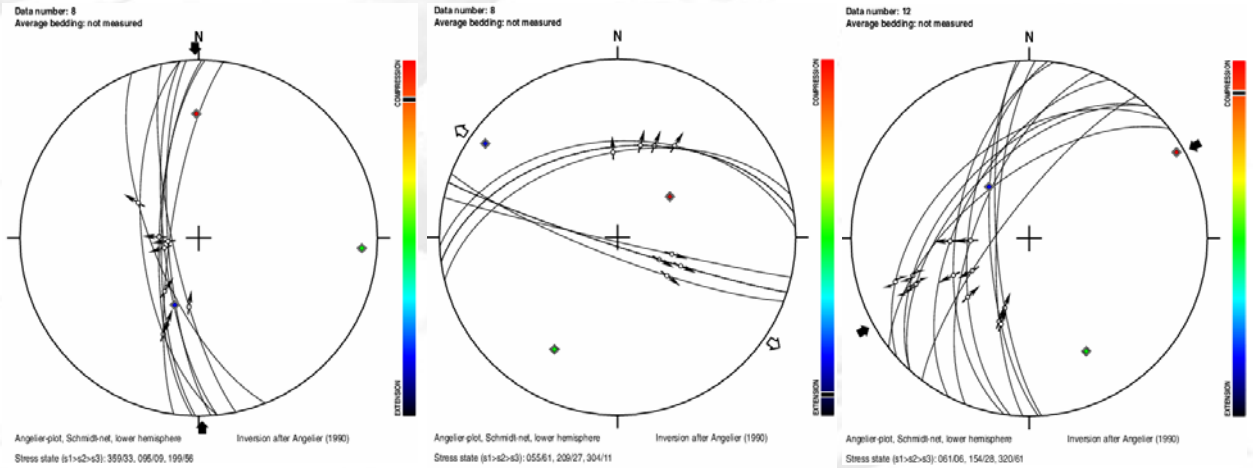
- Before processing, it is possible to sort the data in different groups (using **group codes** in input *.rgf file) and evaluate them **separately** (see Chapter 2)
- If group processing option is **not** is use, data from different groups will displayed and processed **together**
- Using groups setting option, groups will be processed and displayed **group by group**

⬅ Data set with three different groups *not using groups processing* option – all data will be processed together



Same data set processed *using processing option* – all of three data groups defined in input file will be processed and displayed one by one⬇

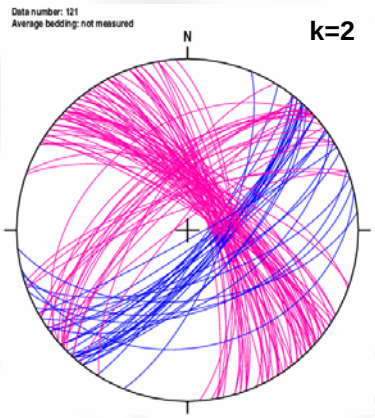
USER DEFINED GROUPS **USING** ➡



⬆ Data set with three different groups *using groups processing* option – all data will be processed on a separate way. With this version, you can display these groups on the same stereonet together

PROCESSING AND DISPLAY

K-MEANS CLUSTERING

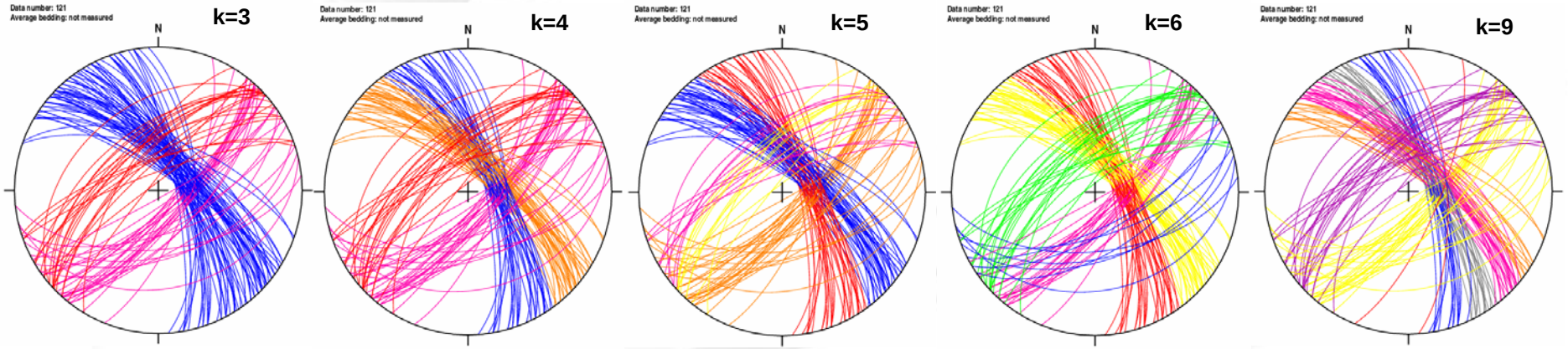


- Built in k-means clustering module is coded to identify **different clusters** of an inhomogeneous data set
- Normal k-means clustering sorts all of data into **k user defied groups**

LIMITATIONS

- First step of k-means clustering is to **sort** the input data in k groups using different algorithms. Please note clustering result depends on this initial step, and sometimes the iteration algorithm produces a mathematically correct solution, but it is just a **local** and not global solution
- User defined k value is the maximum of possible group numbers; because the iteration methodology uses distance calculation, some groups can became **empty** (all data fits better to other groups). If data group becomes empty, groups number was overestimated

Clustering of the same data set into two, three, four five, six and nine clusters with 32.16%, 1.29%, 0.73%, 0.68%, 0.20% and 0.07% relative error. Last solution with nine clusters and 0.07% relative misfit is the optimum solution



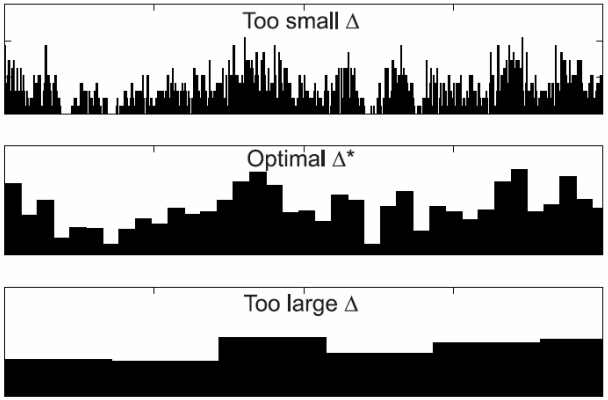
PROCESSING AND DISPLAY

CLUSTERING USING RUP AND ANG VALUES

- This clustering methodology is based on Angelier's (1990) work; according to that, **single phase** slickenside data set has **unimodal** ANG and/or RUP distribution, and **mixed data** set has **bimodal** or **polymodal** ANG and/or RUP distribution. This assumption is checked for each data set, either for the ANG or for the RUP parameter

- To find the **ideal number** of 'valleys' or 'peaks' (=number of clusters) in the distribution of the ANG or UP values, methodology of Shimazaki and Shinomoto (2007) was implemented

- In this case, the main goal is to find the right **bin size** for a data set distribution or histogram. Small bin size results noise (extra, non-existing peaks will appear), too large bin size hides existing peaks



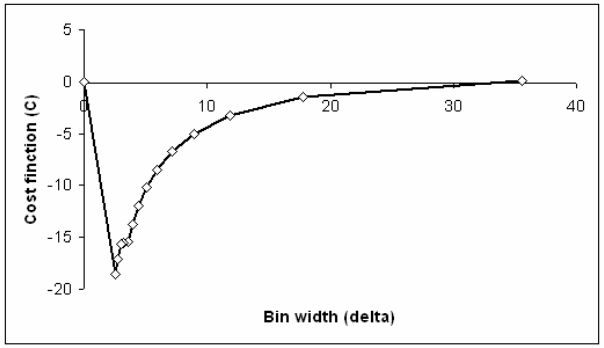
⬆ Histograms for the same data set with too small, optimal, and too large bin size. Figure of Shimazaki and Shinomoto (2007)

METHODOLOGY OF SHIMAZAKI AND SHINOMOTO (2007)

- Divide the data range into **N** bins of width **D**. Count the number of events k_i that enter the i^{th} bin
- Calculate the **mean**, the **variance** and the 'cost function' of the number of events

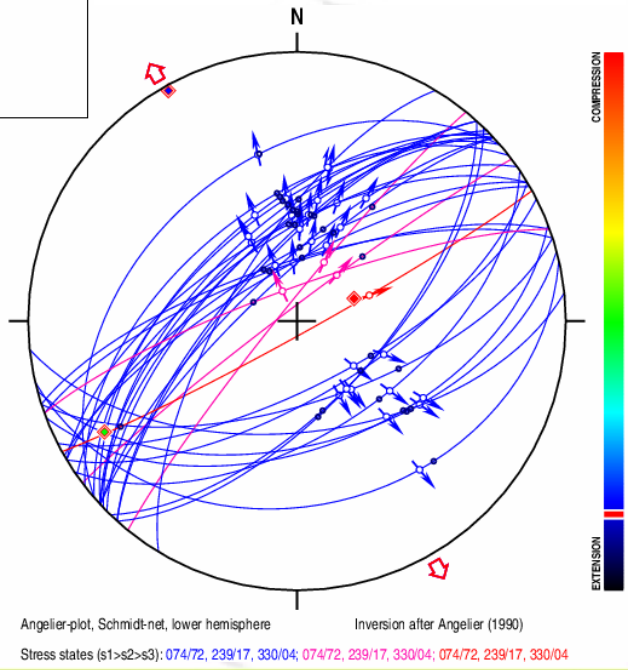
$$k = \frac{1}{N}\sum_{i=1}^{N} k_i, v = \frac{1}{N}\sum_{i=1}^{N}(k_i - k^2), C = \frac{2k - v}{\Delta^2}$$

- Repeat these steps changing **D,** and find **minimum** of cost function 'C'



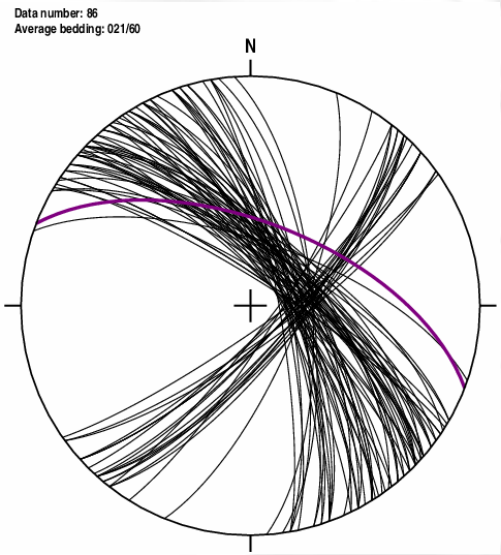
⬅ Bin size versus cost function graph for a test data set; cost function minimum is at ~-18, resulting ~2.5 ideal bin width for clustering

Clustering result of a test data set, based on **RUP** parameter distribution. The input data set could be a heterogeneous set, a mixture of at least two different striae set ➡

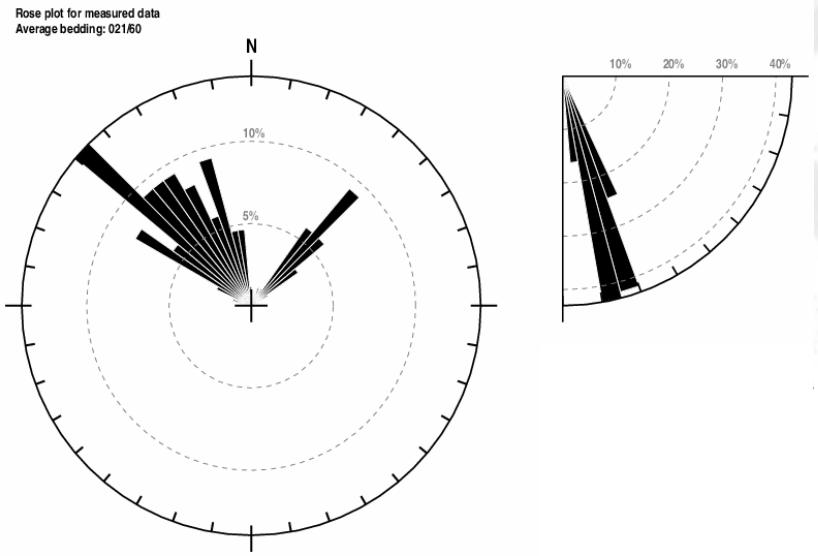


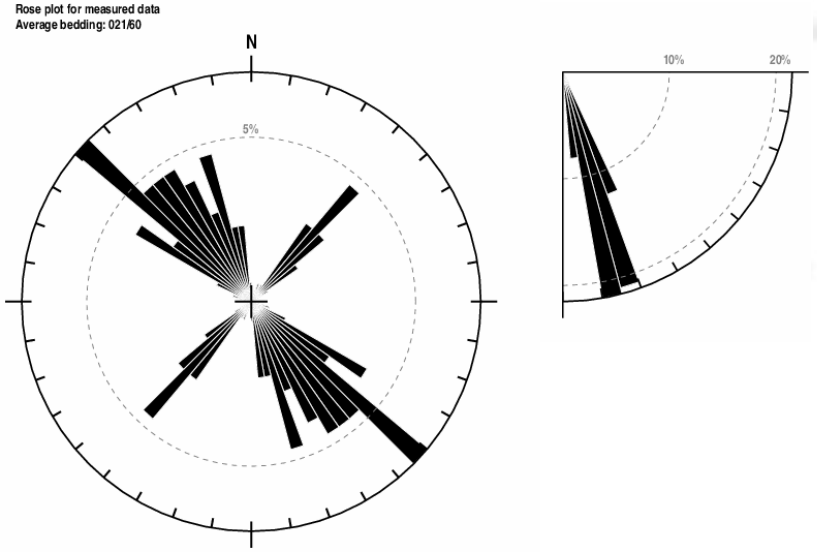
ROSE PLOTS I.

ROSE DIAGRAM TYPE



Same fracture data set plotted on asymmetrical (unidirectional, above) and symmetrical (bidirectional, below) rose diagrams using 5 deg bin sizes ➡

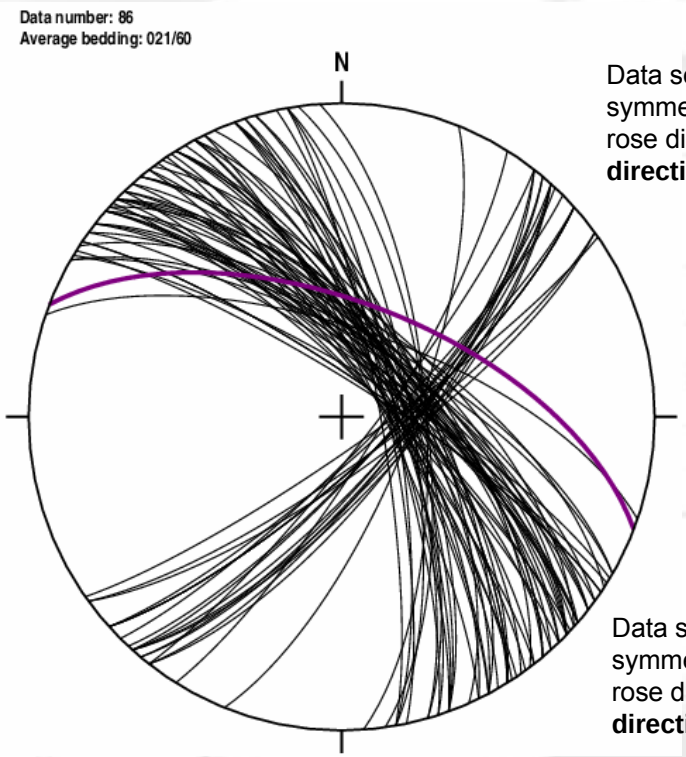




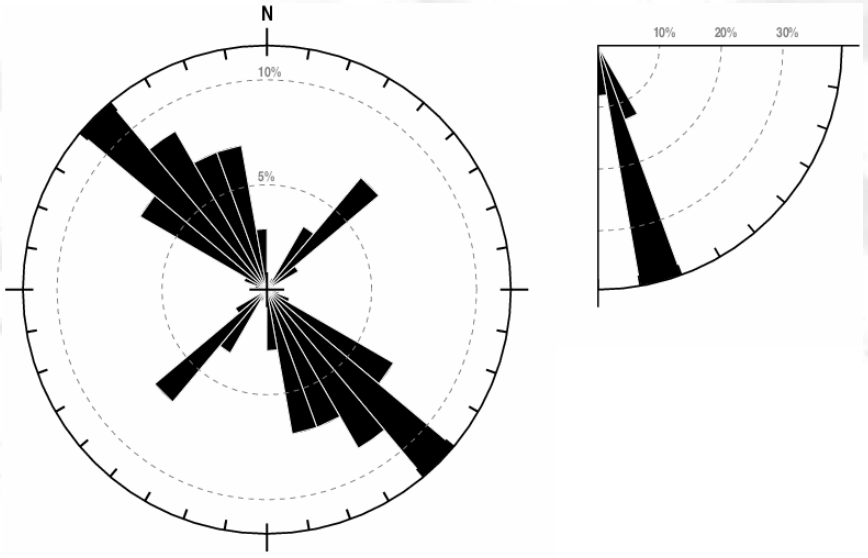
- Rose plots are presenting the **horizontal** and **vertical** distribution of data sets
- Both asymmetric (unidirectional) and symmetric (bidirectional) plots are available

ROSE PLOTS II.

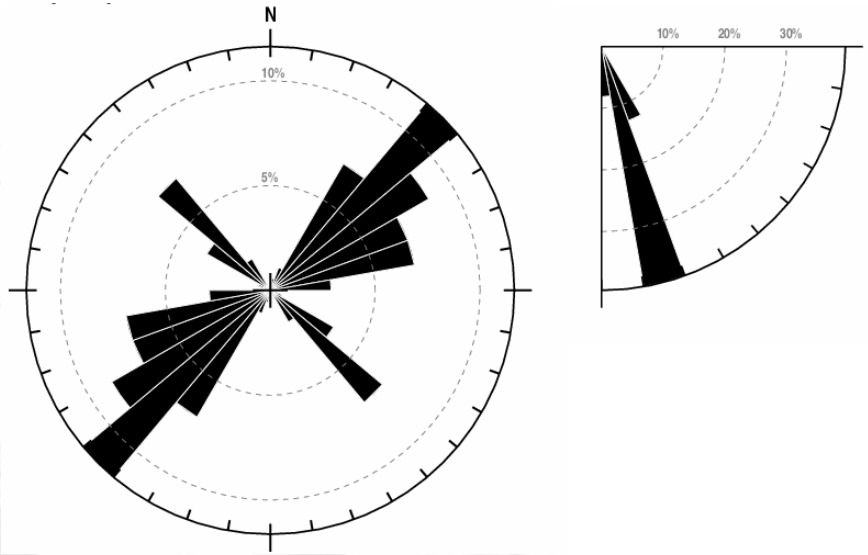
ROSE PLOT DIRECTION



Data set plotted on symmetrical (bidirectional) rose diagrams using **strike direction** of the input data ➡

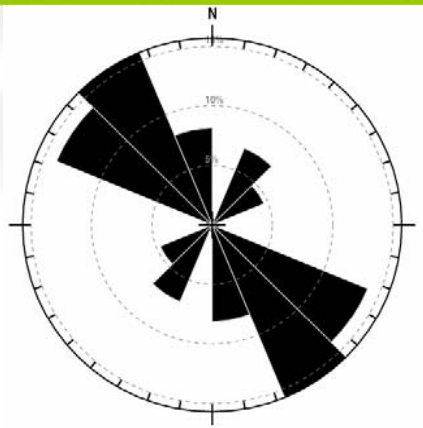


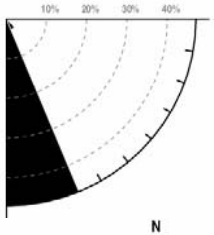
Data set plotted on symmetrical (bidirectional) rose diagrams using **dip direction** of the input data ➡

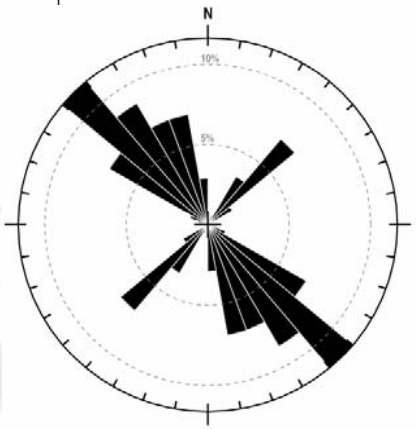


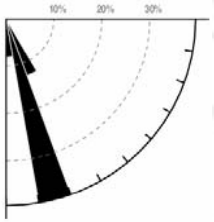
ROSE PLOTS III.

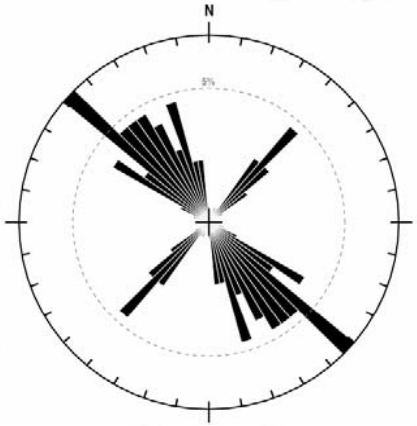
BIN SIZE

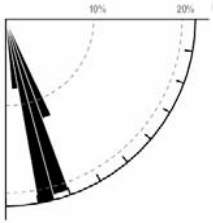


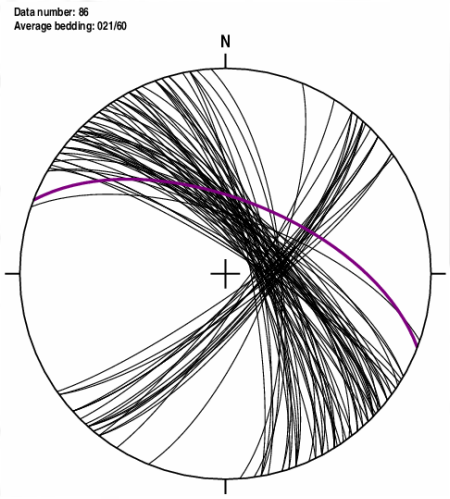


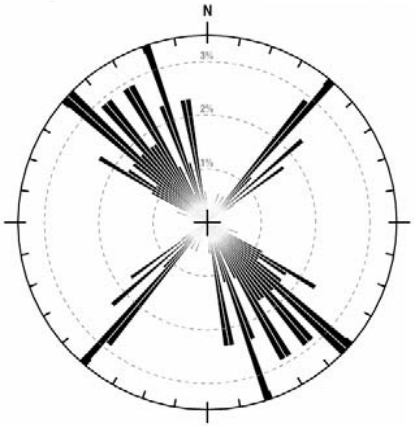


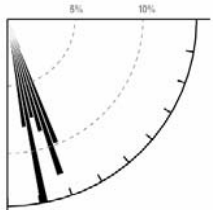












- Rose plots are presenting the **horizontal** and **vertical** distribution of data sets
- Bin size can be 2.5, 4.0, 10.0 and 22.5 degrees

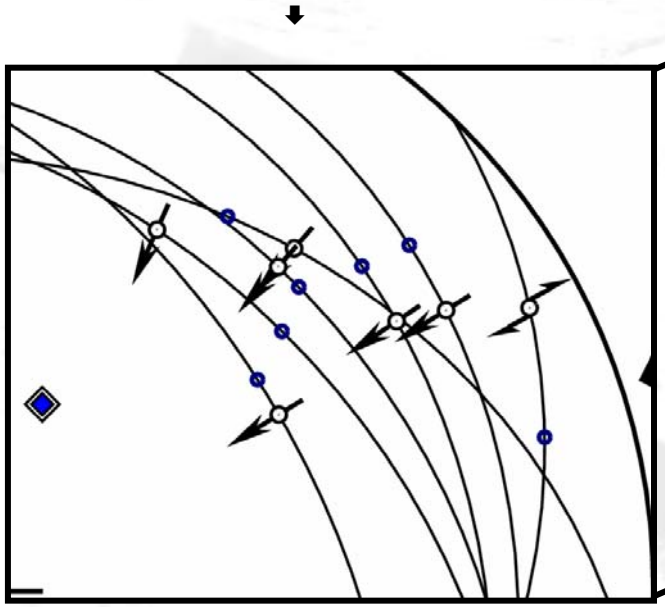
⬆ Same fracture data set plotted on symmetric (bidirectional) rose diagrams using 22.5, 10, 5 and 2.5 deg bin sizes

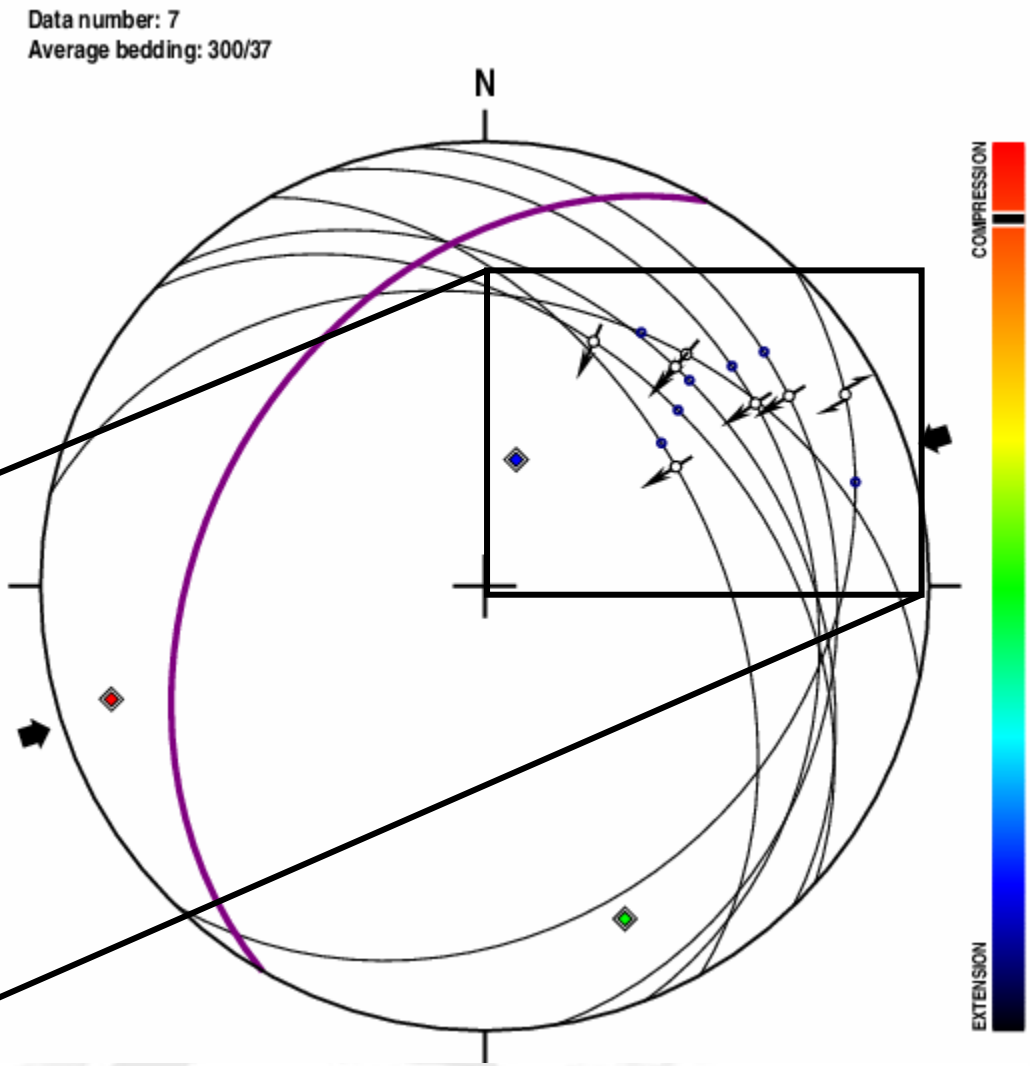
PROCESSING AND DISPLAY

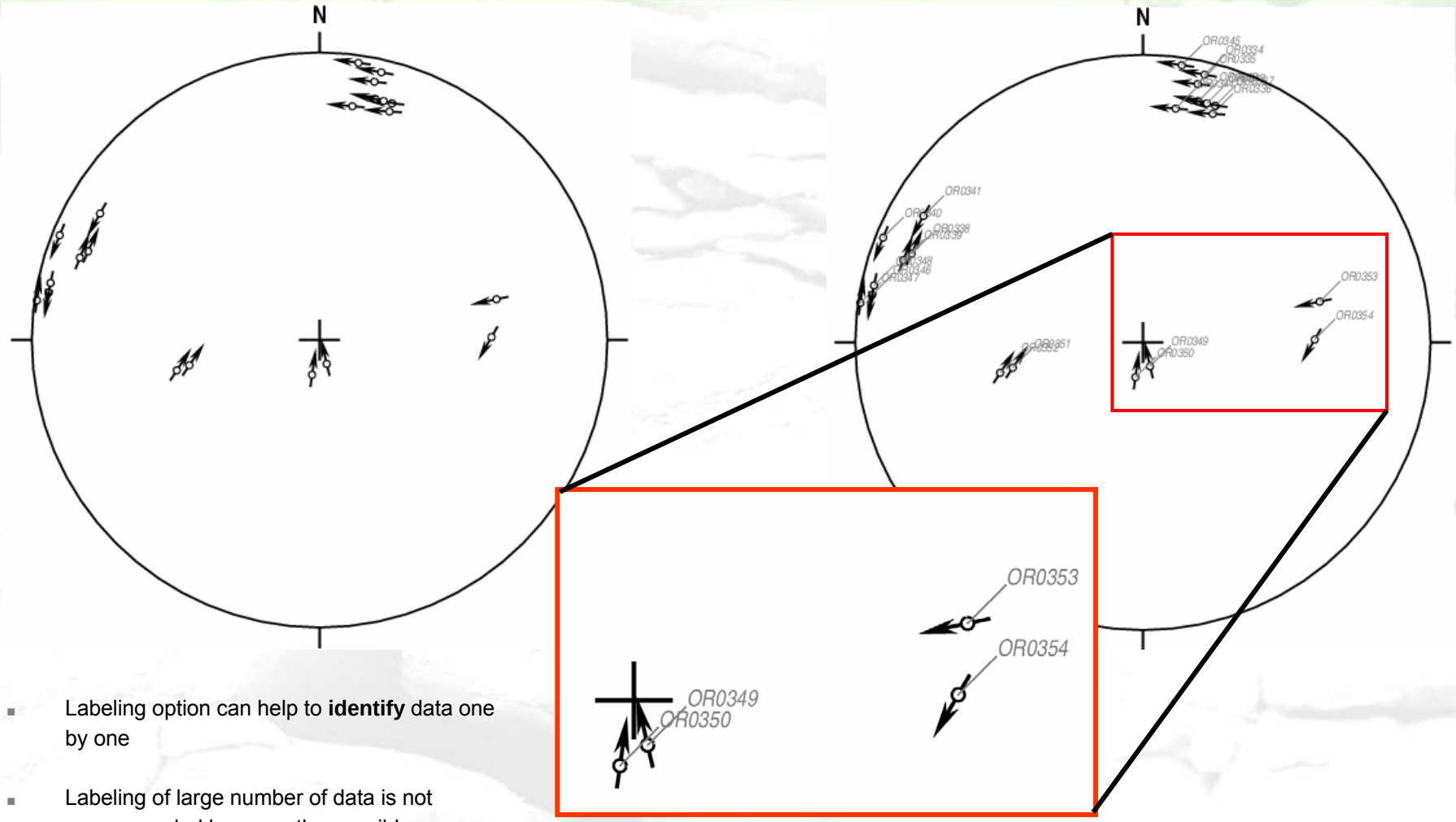
RESOLVED SHEAR DIRECTION DISPLAY

- Using stress inversion results, **ideal shear stress direction** is calculated and plotted
- Small blue circles resolved (ideal) **movement direction** (calculated stress direction) on each planes

Resolved shear directions; small blue circles show the ideal movement direction on each fault planes in the calculated stress field





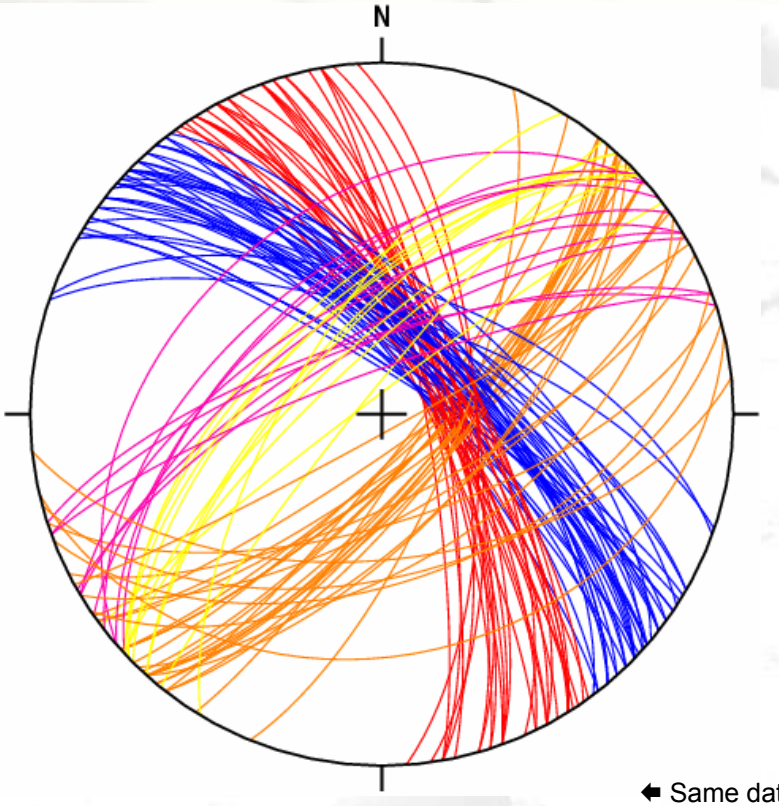


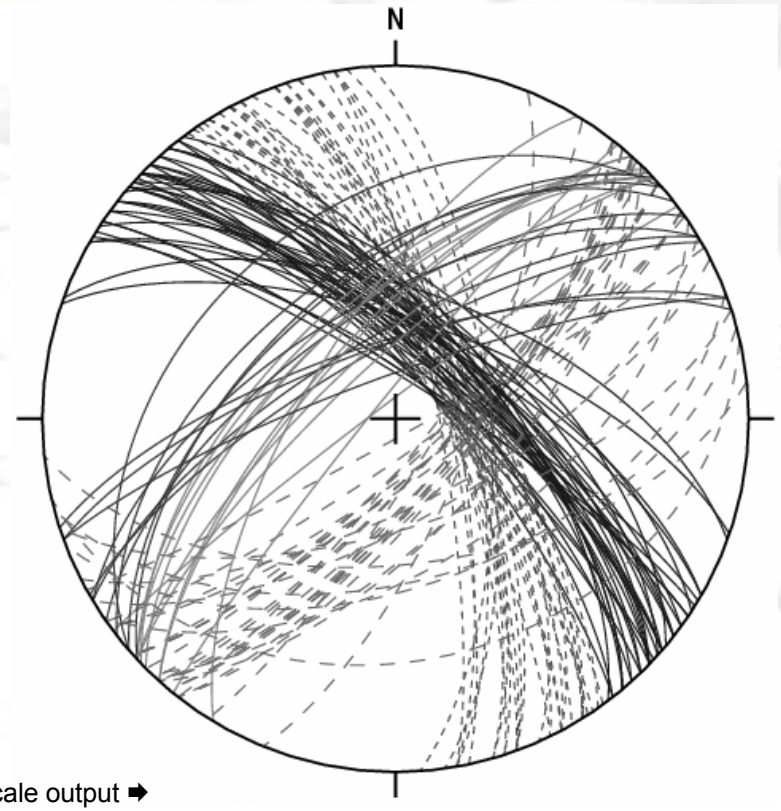
- Labeling option can help to **identify** data one by one
- Labeling of large number of data is not recommended because the possible **overlapping**

Same data set plotted on Hoeppner plot without and with labels (left and right, respectively)

PROCESSING AND DISPLAY

GRAYSCALE IMAGE





⬅ Same data set with RGB and grayscale output ➡

- Grayscale option is ideal for **publications** or black and white printouts
- Different **dash** patterns and three different **grayscale colors** are used to differentiate data using group code or color code

BACKGROUND

COORDINATE SYSTEM AND VARIABLES I.

D, N AND S VECTORS

- Coordinate system for all data processing is: **X** axis points to the **East**, **Y** points to the **North**, and **Z** points **upwards**. All coordinate axes are unit vectors
- All planes, lineations, striated planes and s planes of s-c sets are described using three unit vectors, perpendicular to each other
- In the case of a planar element, N_i unit vector is the **plane normal**, perpendicular to the plane, with Z coordinate ≤ 0 (in the case of overturned beds as well)

$$N = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix} = \begin{bmatrix} \sin(\text{dipdir}) \cdot \sin(\text{dip}) \\ \cos(\text{dipdir}) \cdot \sin(\text{dip}) \\ \cos(\text{dip}) \end{bmatrix}$$

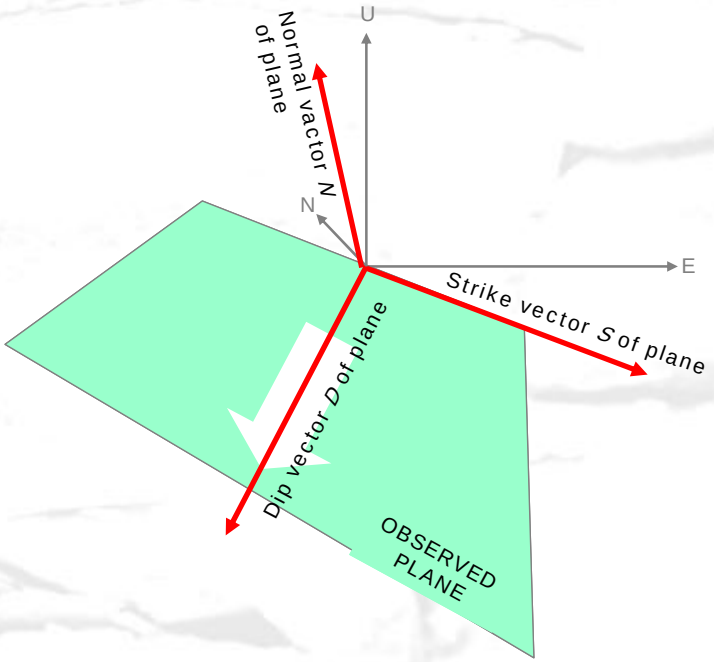
- D_i vector (dip vector) is a unit vector **parallel to the dip of the plane**, with Z coordinate ≥ 0 (in the case of overturned beds as well). Computation methodology of dip vector from dip direction and dip angle data is the following:

$$D = \begin{bmatrix} DX \\ DY \\ DZ \end{bmatrix} = \begin{bmatrix} \sin(\text{dipdir}) \cdot \cos(\text{dip}) \\ \cos(\text{dipdir}) \cdot \cos(\text{dip}) \\ -\sin(\text{dip}) \end{bmatrix}$$

- Calculation of S_i **strike** unit vector (with Z=0) is the following:

$$S = N \otimes D$$

- The methodology is exactly the same for planes holding slickensides, and for s planes of s-c sets
- In the case of **linear** elements, D_i vector points in the dip direction of the linear element. N_i vector is perpendicular to it (as a normal vector of a virtual plane of which D vector is the dip vector), and S_i vector is perpendicular to both of them. Computation methodology is the same as for the planar elements
- In the case of **absolutely symmetric data set**, during eigenvalue computation division by zero can occur. To avoid this, small artificial error ($\pm 10^{-6}$) is added randomly to X, Y or Z coordinates of each plane normal vectors



COORDINATE SYSTEM AND VARIABLES II.

DC, NC AND SC AND VECTORS

- DC_i, NC_i, and SC_i and vectors are used to describe **slickensides** and **s-c** data sets
- In the case of s-c data sets, DC_i, NC_i and SC_i vectors are referring to the dip, normal and strike unit vectors of c plane; in the case of a slickenside, these vectors are referring to the slickenside as a **lineation**
- DC_i vector (dip vector) is a unit vector **parallel to the dip of c plane** or **parallel to slickenside direction**, with Z coordinate ≥ 0. Computation methodology of dip vector from dip direction and dip angle data is the following:

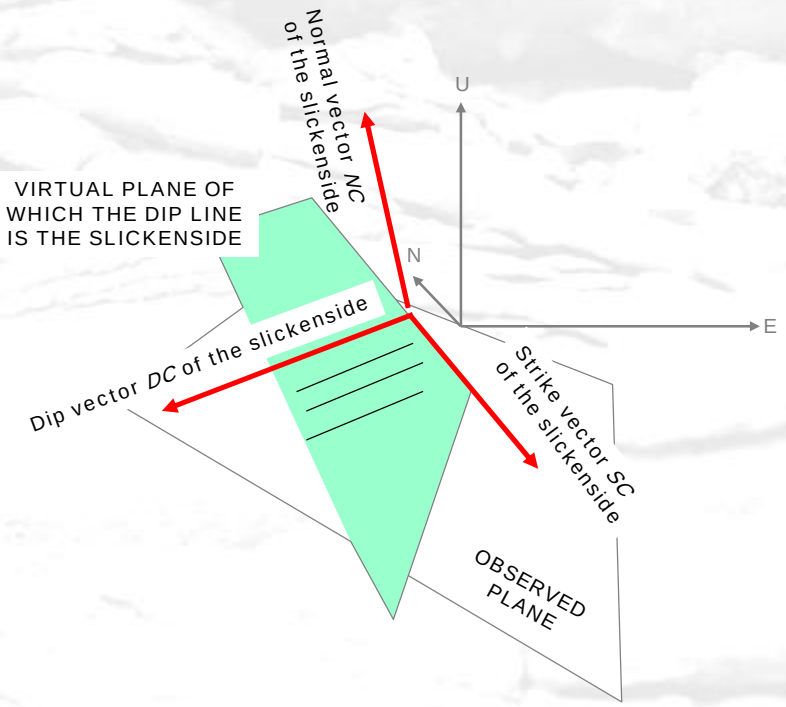
$$DC_i = \begin{bmatrix} DCX \\ DCY \\ DCZ \end{bmatrix} = \begin{bmatrix} \sin(dipdir) \cdot \cos(dip) \\ \cos(dipdir) \cdot \cos(dip) \\ -\sin(dip) \end{bmatrix}$$

- NC_i unit vector is the **c plane normal**, or – in the case of slickensides – **normal vector of a plane of which DC_i vector the dip vector is**, with Z coordinate ≤ 0. Computation methodology of dip vector from dip direction and dip angle data is the following:

$$NC_i = \begin{bmatrix} NCX \\ NCY \\ NCZ \end{bmatrix} = \begin{bmatrix} \sin(dipdir) \cdot \sin(dip) \\ \cos(dipdir) \cdot \sin(dip) \\ \cos(dip) \end{bmatrix}$$

- Calculation of **strike** unit vector (with Z=0) is the following:

$$SC = NC \otimes DC$$



- **Dip vector** is used to display lineations and planes on the stereographic projection:

$$D_i = \begin{bmatrix} DX \\ DY \\ DZ \end{bmatrix}$$

- All lineations are described by **x and y coordinates** on the stereographic projection; these point are calculated from the dip vector for Schmidt- and Wulff-net

- Planes are either **arcs** (in the case of Wulff-net) or **Bezier curves** (in the case Schmidt-projection) using three points to plot them:
  - **dip** of the plane (D_i vector, ❶), and
  - two **strike points** on the stereonet circle (S_i and –S vectors, ❷)

LOWER HEMISPHERE

- On stereographic projection, **x axis** (with unit length) points to **East**, and **y axis** (with unit length) points to the **North**

UPPER HEMISPHERE

- Coordinates above are in lower hemisphere; to use then in upper hemisphere projection, calculate expressions below using the **negative DX** and **DY** values of the dip vector:

$$D_i = \begin{bmatrix} -DX \\ -DY \\ DZ \end{bmatrix}$$

SCHMIDT-NET

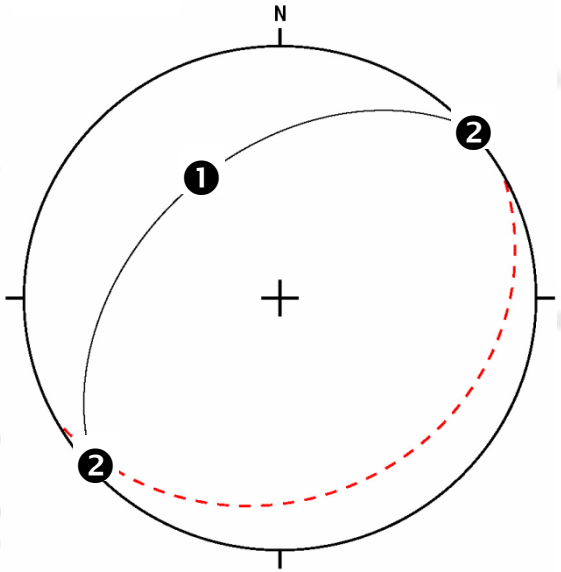
- To calculate x and y coordinates on the Schmidt stereonet , use the following equations:

$$x = \frac{DX}{1.0 - DZ};$$
$$y = \frac{DY}{1.0 - DZ}$$

WULFF-NET

- Wulff net point coordinates are the following:

$$x = \frac{DX}{\sqrt{1.0 - DZ}};$$
$$y = \frac{DY}{\sqrt{1.0 - DZ}}$$



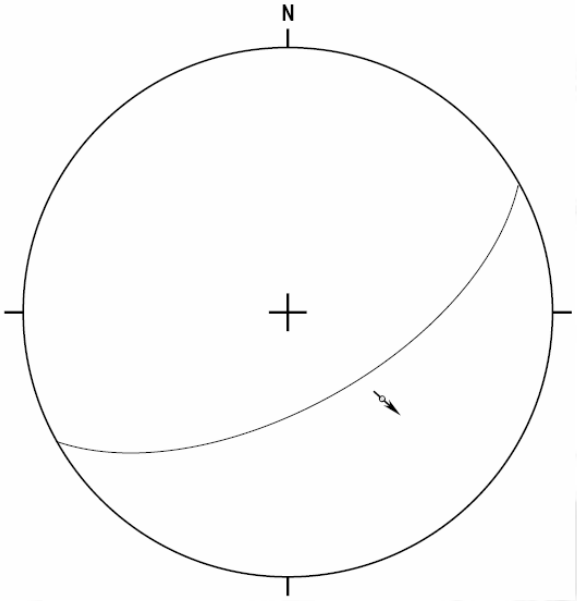
STRIAE MISFIT AND ITS CORRECTION I.

MISFIT ANGLE

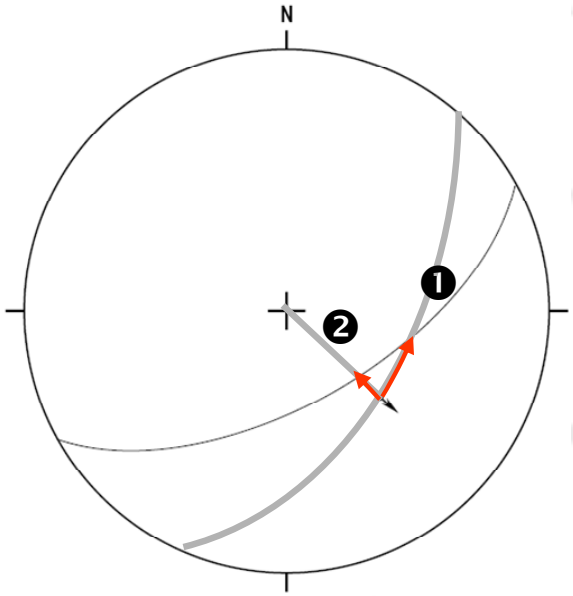
- In ideal case, the measured slickenside vector is on the fault plane surface, and dip vector DC_i of slickenside is **perpendicular** to the plane normal vector N_i
- In the case of measurement as ‘pitch’, slickenside lineation is **definitely** on the fault plane because measurement methodology. In the case of ‘lineation’ observation, a plane and a lineation is measured, which are **not** definitely fitting to each other
- To have the misfit angle between the plane normal N_i and the dip line DC_i, calculate

$$\alpha_i = \left|\arcsin(N_i \cdot DC_i)\right|$$

- It is required to correct ‘lineation’ measurements, but not ‘pitch’ observations. One of the possible correction is ❶❷ to accept dip angle to be correct, and calculate a **corrected dip direction**, or ❷ accept dip direction to be correct, and calculate a **corrected dip angle**



Original data set with misfit ⬆



Possible striae misfit corrections ⬆

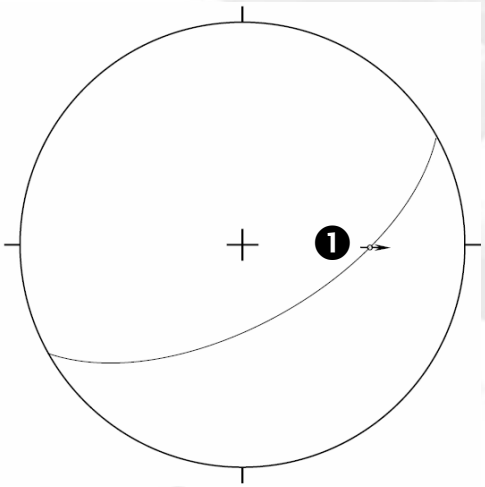
STRIAE MISFIT AND CORRECTION II.

DIP DIRECTION AND DIP CORRECTION, OFFSET CORRECTION

- The first possible correction is to accept dip angle to be correct, and calculate a new **dip direction** fitting on the measured plane

- In the reality, both N_i and SC_i vectors are **perpendicular** to the striae dip vector DC_i, but not necessarily perpendicular to each other. In an ideal case (ie. corrected striae), all of these vectors are perpendicular to each other. So to get the ideal movement direction, calculate the unit vector

$$DC_i = SC_i \otimes N_i$$

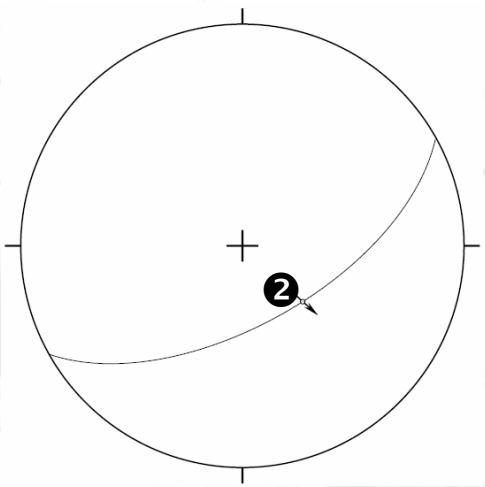


Result of dip direction correction

- In the case of dip angle correction, one accepts dip direction to be correct, and calculate a corrected **dip angle**

- In this case, we are searching for the **intersection** line of two planes: one is the fault plane, and the other one is a virtual plane characterized by NC_i, DC_i and SC_i vectors. The observed slickenside is in the plane, and in the ideal case, it is perpendicular to the plane normal N_i, so it is in the fault plane as well. The only lineation satisfying these two criteria is the intersection line of these planes. To calculate this intersection line, calculate new unit vector

$$DC_i = NC_i \otimes N_i$$



Result of dip angle correction

- Despite several words are describing movement direction (offset) along a slickenside lineation, just two movement directions are visible in reality: the missing block has been moved **upwards** or **downwards** along the striae lineation

- To do the correction of the user defined offset, first the pitch angle (angle between strike and the slickenside direction, 0.0 – 90.0 degrees) will be computed

- During slickenside data processing, user input offset will be changed to be compatible with the followings:
  - if plane dip angle is less than 15 degrees, just ‘normal’ or ‘inverse’ offsets are used
  - if plane dip angle is more than 15 degrees:
    - if the pitch angle is less than 45 degrees, ‘sinistral’ or ‘dextral’ can be used
    - if the pitch angle is more than 45 degrees, ‘normal’ or ‘inverse’ is used

K-MEANS CLUSTERING

- At the initial step, **k data centers** (centroids) are defined; **three dimensional** centroids are used for lineation and plane data, and **six dimensional** ones for striae and s-c data sets. Centroids are unit vectors:

$$C_k = [C_X \quad C_Y \quad C_Z],$$
$$C_k = [C_U \quad C_V \quad C_W \quad C_X \quad C_Y \quad C_Z]$$

- Each plane and lineation data is described by a **three dimensional unit vector**, and each slickenside and s-c data is stored as **six dimensional unit vector:**

$$D_i = [NX \quad NY \quad NZ],$$
$$D_i = [NX \quad NY \quad NZ \quad NCX \quad NCY \quad NCZ]$$

- Input data will be **sorted into k groups** using
  - random grouping, or
  - one by one grouping: first data into the first group, second to the second, etc

1. Distance of each **data to** each **centroid** is calculated on the following way:

$$\delta_i = \left((C_X - NX)^2 + (C_Y - NY)^2 + (C_Z - NZ)^2\right)^2,$$
$$\delta_i = \left((C_U - NX)^2 + (C_V - NY)^2 + (C_W - NZ)^2 + (C_X - NCX)^2 + (C_Y - NCY)^2 + (C_Z - NCZ)^2\right)^2$$

2. Each data is related to the **closest** centroid and became **member of centroid's group**

3. **Calculate new centroid** unit vector from centroid's data group average (n is the clustered database size):

$$C_k = \frac{1}{n}\sum_{i=1}^{n} D_i$$

- Repeat last three steps above either until a **convergence limit**, or user defined **m times**. Please note that in the latter case the iteration can yield suboptimal results.

- Error of clustering is described by **relative error** (ERR) value:

$$\mathrm{ERR} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \cdot 100\%$$

BINGHAM STATISTICS

- **Bingham** directional statistics is used to display directional distribution of fractures

- During the Bingham statistics calculation, a three dimensional '**density ellipsoid**' is generated; the ellipsoid axes show the maximum, intermediate and minimum density direction, and ellipsoid axes length are referring to the geometry of the distribution

- The methodology is a three dimensional **regression**, and the background is almost the same as seen in the case of **PTN** and **NDA** inversion methods (see Chapter 6.7.1 and 6.7.2)

- First generate three unit vectors parallel to the **coordinate axes**:

$$n = \begin{bmatrix} 0.0 \\ 1.0 \\ 0.0 \end{bmatrix}, e = \begin{bmatrix} 1.0 \\ 0.0 \\ 0.0 \end{bmatrix}, u = \begin{bmatrix} 0.0 \\ 0.0 \\ 1.0 \end{bmatrix},$$

- Calculate the following **directional tensor** for each i[th] data set (ang is the angle between two vectors):

$$E_i = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix},$$

$$\begin{aligned} \varepsilon_{11} &= ang(N,e) \cdot ang(P,e), \\ \varepsilon_{12} &= ang(N,e) \cdot ang(P,n), \\ \varepsilon_{13} &= ang(N,e) \cdot ang(P,u), \\ \varepsilon_{22} &= ang(N,n) \cdot ang(P,n), \\ \varepsilon_{23} &= ang(N,n) \cdot ang(P,u), \\ \varepsilon_{33} &= ang(N,u) \cdot ang(P,u) \end{aligned}$$

- Calculate the following sum for the entire data set:

$$E = \sum_{i=1}^{n} E_i,$$

- **Eigenvectors** of matrice E points to maximum, minimum and intermediate axes if the directional distribution ellipsoid

- **Eigenvalues** of matrix E can show the geometry of fracture distribution ellipsoid

AVERAGE CALCULATION AND SINGULARITY

- In the case of a **single** data set, the data set itself is the average

- If two data forms the entire data set, **mathematical average** of their N vectors is the average. If N_1 is the opposite of N_2, no average is calculated (because it would be a null vector)

- **Bingham** directional statistics is used to calculate data average for data sets with more than 2 records. If the input data set is symmetrical, the final directional tensor will be singular, with a really small determinant more or less equal to zero

$$E_i = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix} = 0$$

- During eigenvalue and eigenvector calculation, really small (almost zero) determinant can result in '**division by zero**' or in division by a really small number (generating unreliable result)

```
33/70 normal
34/71 normal
35/72 normal
33/70 overturned
34/71 overturned
35/72 overturned
```

Singular data set example – processing is not possible and the user is informed on that

- To avoid this problem, during data average calculation for each data sets with more than two records, directional tensor is calculated prior the eigenvalue and eigenvector computation

- If the determinant of the directional tensor is less than 10^{-25}, no average is calculated, and the user is informed on the singularity of the input data set

- This singularity check makes sense from structural geological point as well: what is the average bedding of a 30/70 normal and 30/70 overturned bedding measurement?

INVERSION METHODOLOGIES I.

REGRESSION AFTER SPANG (1972)

- Define angle a between the **movement direction** (slip vector) and the **maximum stress direction**

- Declare North, East and Upwards **unit vectors** below:

$$n_i = \begin{bmatrix} 0.0 \\ 1.0 \\ 0.0 \end{bmatrix}, e_i = \begin{bmatrix} 1.0 \\ 0.0 \\ 0.0 \end{bmatrix}, u_i = \begin{bmatrix} 0.0 \\ 0.0 \\ 1.0 \end{bmatrix},$$

- Calculate p_i and t_i vectors for each data set, referring to **compression** and **tension** directions:

$$p_i = \begin{bmatrix} (1-\alpha)DCX - \alpha NX \\ (1-\alpha)DCY - \alpha NY \\ (1-\alpha)DCZ - \alpha NZ \end{bmatrix},$$

$$t_i = \begin{bmatrix} (1-\alpha)NX + \alpha DCX \\ (1-\alpha)NY + \alpha DCY \\ (1-\alpha)NZ + \alpha DCZ \end{bmatrix},$$

- Calculate the following **sums**:

$$P = \sum_{i=1}^{n} p_i,$$

$$T = \sum_{i=1}^{n} t_i,$$

- The stress tensor **T** is the following:

$$T_i = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix},$$

$$\sigma_{11} = ang(P,e) \cdot ang(P,e) - ang(T,e) \cdot ang(T,e),$$
$$\sigma_{12} = ang(P,e) \cdot ang(P,n) - ang(T,e) \cdot ang(T,n),$$
$$\sigma_{13} = ang(P,e) \cdot ang(P,u) - ang(T,e) \cdot ang(T,u),$$
$$\sigma_{22} = ang(P,n) \cdot ang(P,n) - ang(T,n) \cdot ang(T,n),$$
$$\sigma_{23} = ang(P,n) \cdot ang(P,u) - ang(T,n) \cdot ang(T,u),$$
$$\sigma_{33} = ang(P,u) \cdot ang(P,u) - ang(T,u) \cdot ang(T,u)$$

INVERSION METHODOLOGIES II.

REGRESSION AFTER TURNER (1953)

- Define angle α between the **movement** and the **maximum stress** direction
- Define North, East and Upwards **unit direction vectors** below:

$$n = \begin{bmatrix} 0.0 \\ 1.0 \\ 0.0 \end{bmatrix}, e = \begin{bmatrix} 1.0 \\ 0.0 \\ 0.0 \end{bmatrix}, u = \begin{bmatrix} 0.0 \\ 0.0 \\ 1.0 \end{bmatrix},$$

- Calculate **pressure** (P), **tension** (T) and **neutral** (N) vectors for eash data:

$$p_i = \begin{bmatrix} (1-\alpha)DCX - \alpha NX \\ (1-\alpha)DCY - \alpha NY \\ (1-\alpha)DCZ - \alpha NZ \end{bmatrix}, t_i = \begin{bmatrix} (1-\alpha)NX + \alpha DCX \\ (1-\alpha)NY + \alpha DCY \\ (1-\alpha)NZ + \alpha DCZ \end{bmatrix}, n_i = p_i \otimes t_i$$

- Calculate P_i, T_i and N_i tensors for each data set referring to **compression**, **tension** and **neutral** directions:

$$P_i = \begin{bmatrix} ang(p_i,e)\cdot ang(p_i,e) & ang(p_i,e)\cdot ang(p_i,n) & ang(p_i,e)\cdot ang(p_i,u) \\ ang(p_i,e)\cdot ang(p_i,n) & ang(p_i,n)\cdot ang(p_i,n) & ang(p_i,n)\cdot ang(p_i,n) \\ ang(p_i,e)\cdot ang(p_i,u) & ang(p_i,n)\cdot ang(p_i,n) & ang(p_i,u)\cdot ang(p_i,u) \end{bmatrix},$$

$$T_i = \begin{bmatrix} ang(t_i,e)\cdot ang(t_i,e) & ang(t_i,e)\cdot ang(t_i,n) & ang(t_i,e)\cdot ang(t_i,u) \\ ang(t_i,e)\cdot ang(t_i,n) & ang(t_i,n)\cdot ang(t_i,n) & ang(t_i,n)\cdot ang(t_i,n) \\ ang(t_i,e)\cdot ang(t_i,u) & ang(t_i,n)\cdot ang(t_i,n) & ang(t_i,u)\cdot ang(t_i,u) \end{bmatrix},$$

$$N_i = \begin{bmatrix} ang(n_i,e)\cdot ang(n_i,e) & ang(n_i,e)\cdot ang(n_i,n) & ang(n_i,e)\cdot ang(n_i,u) \\ ang(n_i,e)\cdot ang(n_i,n) & ang(n_i,n)\cdot ang(n_i,n) & ang(n_i,n)\cdot ang(n_i,n) \\ ang(n_i,e)\cdot ang(n_i,u) & ang(n_i,n)\cdot ang(n_i,n) & ang(n_i,u)\cdot ang(n_i,u) \end{bmatrix}$$

- Calculate **sum** tensors

$$P = \sum_{i=1}^{n} P_i,$$

$$T = \sum_{i=1}^{n} T_i,$$

$$N = \sum_{i=1}^{n} N_i$$

- Calculate the **eigenvalues and eigenvectors** for all of **three** tensors:

$$EVAL_P = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}, EVEC_P = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix},$$

$$EVAL_T = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}, EVEC_T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

$$EVAL_N = \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix}, EVEC_N = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

- The **reduced stress tensor** will be:

$$\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} = \begin{bmatrix} P_1 & 0 & 0 \\ 0 & N_2 & 0 \\ 0 & 0 & T_3 \end{bmatrix}$$

- The maximum, intermediate and minimum stress directions will be $EVEC_P$, $EVEC_T$, and $EVEC_N$, respectively

INVERSION METHODOLOGIES III.

REGRESSION AFTER FRY (1999)

- Use Angelier's (1979) assumption for each slickenside data: $\boldsymbol{b_i \cdot T \cdot n_i = 0}$, where

$$b_i = DC \otimes N = [b_1 \quad b_2 \quad b_3],$$

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix},$$

$$n_i = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

- In this case

$$[b_1 \quad b_2 \quad b_3] \cdot \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} =$$

$$b_1\sigma_{11}n_1 + b_1\sigma_{12}n_2 + b_1\sigma_{13}n_3 +$$
$$b_2\sigma_{12}n_1 + b_2\sigma_{22}n_2 + b_2\sigma_{23}n_3 +$$
$$b_3\sigma_{13}n_1 + b_3\sigma_{23}n_2 + b_3\sigma_{33}n_3 = 0$$

- Use **regression** to search for the best fitting values of T; first compute the coefficients of s_{11}, s_{22}, s_{33}, s_{12}, s_{23} and s_{13}, and they are the values of a **six-dimensional unit vector** p_i:

$$p_i = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix} = \begin{bmatrix} b_1n_1 \\ b_2n_2 \\ b_3n_3 \\ b_1n_2 + b_2n_1 \\ b_2n_3 + b_3n_2 \\ b_1n_3 + b_3n_1 \end{bmatrix}$$

- Calculate the **sum of the inner products** (**second moment tensor**) of p_i vectors:

$$P = \sum_{i=1}^{n} p_i \cdot p_i^T$$

- Find the **eigenvalues and eigenvectors** of this 6x6 symmetrical matrice, e.g. use the Jacobi algorithm

- The eigenvector of the lowest eigenvalue will be:

$$E_{vec} = [0.57735 \quad 0.57735 \quad 0.57735 \quad 0.00000 \quad 0.00000 \quad 0.00000]$$

- Choose the eigenvector of the **second** lowest eigenvalue:

$$E_{vec} = [E_1 \quad E_2 \quad E_3 \quad E_4 \quad E_5 \quad E_6]$$

- These values (**the coefficients of $p_1 - p_6$**) will be the members of the stress tensor:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} E_1 & E_4 & E_6 \\ E_4 & E_2 & E_5 \\ E_6 & E_5 & E_3 \end{bmatrix}$$

INVERSION METHODOLOGIES IV.

REGRESSION AFTER SHAN ET AL. (2003)

- Shan et al's (2003) methodology is really similar to Fry's (1999) solution; the most important difference is the **reduction in the dimension numbers** using basic assumption after Angelier (1979): $s_{11} + s_{22} + s_{33} = 0$

- Calculate the **best fitting stress tensor** for the equation below:

$$b_i \cdot T \cdot n_i = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \cdot \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} =$$

$$b_1\sigma_{11}n_1 + b_1\sigma_{12}n_2 + b_1\sigma_{13}n_3 + b_2\sigma_{12}n_1 + b_2\sigma_{22}n_2 + b_2\sigma_{23}n_3 + b_3\sigma_{13}n_1 + b_3\sigma_{23}n_2 + b_3\sigma_{33}n_3 = 0,$$

$$\sigma_{11} + \sigma_{22} + \sigma_{33} = 0$$

- Use **regression** to search for the best fitting values of the stress tensor: first compute the coefficients of s_{11}, s_{22}, s_{12}, s_{13} and s_{23}; they are the values of a **five-dimensional unit vector** p_i:

$$p_i = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} b_1n_1 - b_3n_3 \\ b_2n_2 - b_3n_3 \\ b_1n_2 + b_2n_1 \\ b_1n_3 + b_3n_1 \\ b_2n_3 + b_3n_2 \end{bmatrix}$$

- Calculate the **sum of the inner products** (**second moment tensor**) of p_i vectors:

$$P = \sum_{i=1}^{n} p_i \cdot p_i^T$$

- Find the **eigenvalues and eigenvectors** of this 5x5 symmetrical matrice, e.g. use Jacobi algorithm

- The eigenvector of the **lowest eigenvalue** will be:

$$E_{vec} = \begin{bmatrix} E_1 & E_2 & E_3 & E_4 & E_5 \end{bmatrix}$$

- These values (the coefficients of p_1 – p_6) will be the **members of the stress tensor**:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} E_1 & E_3 & E_4 \\ E_3 & E_2 & E_5 \\ E_4 & E_5 & 0 - E_1 - E_2 \end{bmatrix}$$

INVERSION METHODOLOGIES V.

REGRESSION AFTER MICHAEL (1984)

- Calculate the following **orientation matrix** using fault plane normal directions for each slickenside:

$$M_i = \begin{bmatrix} n_1 - {n_1}^3 + n_1 {n_3}^2 & n_2 - 2n_2 n_1^2 & n_3 - 2n_3 n_1^2 & -n_1 n_2^2 + n_1 n_3^2 & -2n_1 n_2 n_3 \\ -n_2 n_1^2 + n_2 n_3^2 & n_1 - 2n_1 n_2^2 & -2n_1 n_2 n_3 & n_2 - {n_2}^3 + n_2 {n_3}^2 & n_3 - 2n_3 n_2^2 \\ -n_3 {n_1}^2 - n_3 + {n_3}^3 & -2n_1 n_2 n_3 & n_1 - 2n_1 n_3^2 & -n_2^2 n_3 - n_3 + {n_3}^3 & n_2 - 2n_2 n_3^2 \end{bmatrix},$$

$$\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix}$$

- Calculate the **movement matrix** for each slickenside data:

$$B_i = \begin{bmatrix} DCX \\ DCY \\ DCZ \end{bmatrix}$$

- The matrix form of the **stress tensor** is:

$$X = \begin{bmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{22} \\ \sigma_{23} \end{bmatrix}, \sigma_{11} + \sigma_{22} + \sigma_{33} = 0$$

- Calculate following **hypermatrices** for all of *n* faults using orientation matrices and movement matrices:

$$\mathrm{M} = \begin{bmatrix} \mathrm{M}_1^{\mathrm{T}} \\ \vdots \\ \mathrm{M}_i^{\mathrm{T}} \\ \vdots \\ \mathrm{M}_n^{\mathrm{T}} \end{bmatrix}, \mathrm{B} = \begin{bmatrix} \mathrm{B}_1^{\mathrm{T}} \\ \vdots \\ \mathrm{B}_i^{\mathrm{T}} \\ \vdots \\ \mathrm{B}_n^{\mathrm{T}} \end{bmatrix},$$

- Linear equation system to be solved for *X* is *MX* = *B*

- **Rectangular** (nxn) form of matrice M is needed to solve this equation: calculate matrices $A = M^{\mathrm{T}}M$ *and* $C = M^{\mathrm{T}}B$, and the equation system to be solved will be *AX* = *C*

- To solve this equation system, use e.g. **LU decomposition** methodology to compute members of matrix X; because matrix *A* is symmetrical, simple Gaussian elimination will be not working

INVERSION METHODOLOGIES VI.

REGRESSION AFTER ANGELIER (1990)

- Angelier's (1990) methodology calculates the **misfit vector** υ between the measured (DC_i) and calculated (τ) shear stress vectors; length of DC_i is unknown, and it was assumed to be $\lambda = \sqrt{3}/2$

- Basic assumption after Angelier (1979) is to suppose $s_{11} + s_{22} + s_{33} = 0$ and $s_{11}^2 + s_{22}^2 + s_{33}^2 = 3/2$; the stress tensor below satisfies these criteria:

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} \cos(\psi) & \alpha & \gamma \\ \alpha & \cos\left(\psi + \frac{2\pi}{3}\right) & \beta \\ \gamma & \beta & \cos\left(\psi + \frac{4\pi}{3}\right) \end{bmatrix}$$

- **Length of vector** υ is expressed as a function of the plane **normal vector**, the **slickenside lineation**, the assumed (a priori) shear stress vector length λ and the members of stress tensor (α, β, γ and ψ; see equations A6 and A8 in Angelier 1990)

- To minimize length of υ vector, Angelier (1990) calculates the following **partial derivates** (see equation A10 in Angelier 1990):

$$\frac{1}{2}\frac{\partial v^2}{\partial \alpha} = 0, \frac{1}{2}\frac{\partial v^2}{\partial \beta} = 0, \frac{1}{2}\frac{\partial v^2}{\partial \gamma} = 0, \frac{1}{2}\frac{\partial v^2}{\partial \psi} = 0$$

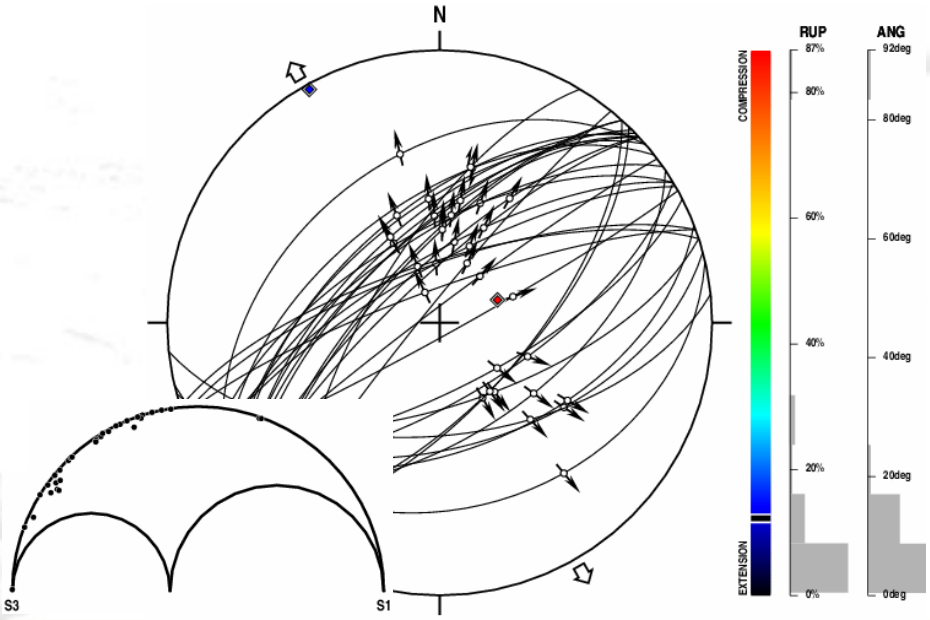
- The final equation (A15, using equation A11 in Angelier 1990) is the function of ψ on the fourth power

- A fourth power equation has 1) four real solutions, 2) two real and two complex solution, or 3) all solutions are zero; **usually**, this equation has **two real solutions** (ψ_1 and ψ_2) and two complex ones

- Using ψ value, express α, β and γ and **generate the stress tensor** (see equation A12 in Angelier 1990)

- To find the best fitting stress tensor, generate two stress tensors using both ψ_1 and ψ_2 values and choose the **best fitting one**

Angelier's (1990) AVB data set processed by Angelier's (1990) methodology ⬇

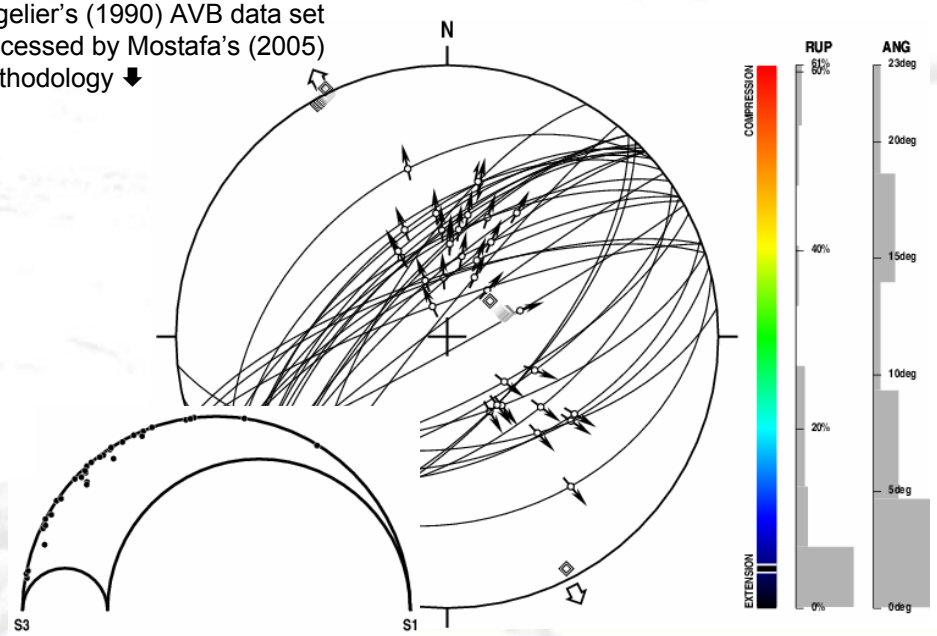


INVERSION METHODOLOGIES VII.

ITERATIVE REGRESSION AFTER MOSTAFA (2005)

- Mostafa's (2005) methodology is an **iteration in shear stress vector length**. Value of λ was assumed by Angelier (1990) to be √3/2
- Mostafa (2005) uses Angelier's (1990) solution to generate the stress field and calculate τ shear vector length
- In the second iteration step, **λ** (measured shear stress vector length) is **replaced by the length of τ** (calculated shear vector) in all of equations, and a new stress tensor calculation is carried out
- In the next steps, values of **τ are re-calculated**; after the 10-15th iteration no major change in stress estimators is visible
- One of the key problems of inversion methodologies is the measured shear vector (DCi) length calculation
- Because basic inversion equations are nonlinear in the sense of shear vector length, either 1) assumption on the original vector length is used (λ=1 in the case of Michael 1984, and λ = √3/2 was used by Angelier 1990), or 2) iteration in the shear vector length
- Using Angelier's 1990 methodology, Mostafa (2005) proposed the following iteration methodology:
  - assume λ = √3/2
  - calculate the stress tensor using Angelier's (1990) methodology
  - calculate the stress vector, the normal stress and than the shear stress vectors
  - calculate shear vector length |t|
  - use λ = |t| for the next iteration step
- Using Michael's (1984) methodology, is it possible to use the same iteration methodology with λ=1 start criteria

Angelier's (1990) AVB data set processed by Mostafa's (2005) methodology ⬇



INVERSION METHODOLOGIES VIII.

BRUTE FORCE INVERSION

GENERAL METHODOLOGY

- Generate a **set of stress tensors**, calculate the **resolved shear** on the given data set, choose the **10 best fitting tensors**, and do a **detailed investigation** in their proximity. Accept the **best fitting** stress tensor

STRESS TENSORS

- Use **Kaalsbeek counting net** to generate a series of σ_3 directions (331 points)
- **Rotate** σ_1 axes around σ_3 axis with 10 degrees steps (18 rotations, 0.0 – 170 degrees)
- **Calculate** σ_2 directions
- Use **0.0 – 1.0 phi values** for the stress tensor geometry with 0.1 steps (11 different geometries)

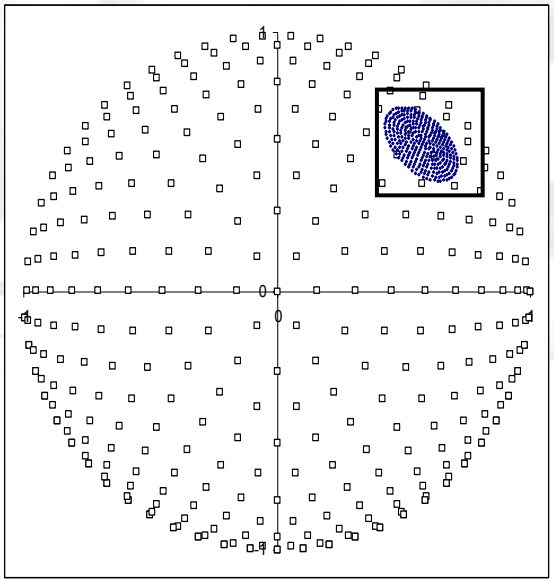
INITIAL ITERATION

- Calculate **resolved shear directions** on all fault planes; do altogether 331 x 18 x 11 = 65538 iterations
- Calculate the **cumulative angular misfit** (ANG) between the observed and the resolved shear directions
- Store stress tensor parameters for the **10 smallest misfit** values

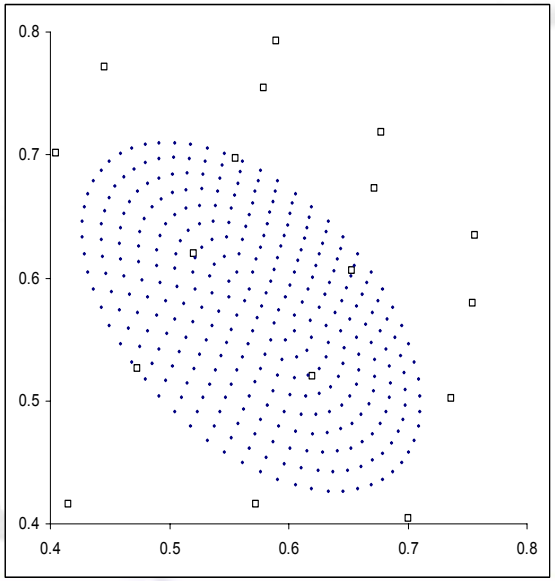
HIGH RESOLUTION ITERATION

- Generate a **detailed 'zoom in' Kaalsbeek net** around each local minima
- **Repeat the iteration** for the local maxima using the high resolution Kaalsbeek net
- Any stress tensor with the **smallest cumulative angular misfit** is the solution of the brute force inversion

➡ Kaalsbeek counting net to generate σ_3 directions



High resolution 'zoom in' area with a high resolution Kaalsbeek counting net ➡



NEGATIVE OF STRESS TENSOR

- In several cases (ie. in the case of Fry 1999 and Shan et al. 2003 methodology) both **stress tensor** and the **negative of the stress tensor**:

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}$$

and

$$T = \begin{bmatrix} -\sigma_{11} & -\sigma_{12} & -\sigma_{13} \\ -\sigma_{12} & -\sigma_{22} & -\sigma_{23} \\ -\sigma_{13} & -\sigma_{23} & -\sigma_{33} \end{bmatrix}$$

could be a good solution after stress field computation

- Because the coordinate conventions are the same for all of methodologies, it must be **checked** for each inversion technologies that the result stress tensor, or its negative is the correct one
- One of the possible solution to find the best fitting tensor, is to **calculate the sum of the angular misfits** ANG (see Chapter 6.9.2) between observed and computed shear direction 1) for the stress tensor and 2) for its negative, and choose the best fitting one

NEGATIVE OF STRESS VECTOR

- Before stress estimator calculation, the right **stress vector orientation** must be checked for each data sets. Because the stress data convention, stress vector must be **positive** if **pointing into the deformed body**. After inverting a homogenous data group, this will be true for the majority of data sets, but not for all of them
- Calculate the stress vector and its magnitude for each slickensides:

$$\sigma = n_i \cdot T = \begin{bmatrix} \sigma_X \\ \sigma_Y \\ \sigma_Z \end{bmatrix} = \begin{bmatrix} n_1\sigma_{11} + n_2\sigma_{12} + n_3\sigma_{13} \\ n_1\sigma_{12} + n_2\sigma_{22} + n_3\sigma_{23} \\ n_1\sigma_{13} + n_2\sigma_{23} + n_3\sigma_{33} \end{bmatrix},$$

$$|\sigma| = n_1\sigma_X + n_2\sigma_Y + n_3\sigma_Z$$

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}, n_i = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

- If the inversion methodology produces a **negative of stress vector** (pointing outside of the deformed body) **and magnitude of stress in positive**, or the **stress vector** is positive **and magnitude of stress in negative**, use the negative of stress vector

STRESS VECTORS AND STRESS ESTIMATORS I.

STRESS VECTOR, NORMAL AND SHEAR STRESS ON PLANE

- To have the **stress vector** effecting the i[th] plane, calculate

$$\sigma = n_i \cdot T = \begin{bmatrix} \sigma_X \\ \sigma_Y \\ \sigma_Z \end{bmatrix} = \begin{bmatrix} n_1\sigma_{11} + n_2\sigma_{12} + n_3\sigma_{13} \\ n_1\sigma_{12} + n_2\sigma_{22} + n_3\sigma_{23} \\ n_1\sigma_{13} + n_2\sigma_{23} + n_3\sigma_{33} \end{bmatrix}$$

$$T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}, n_i = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

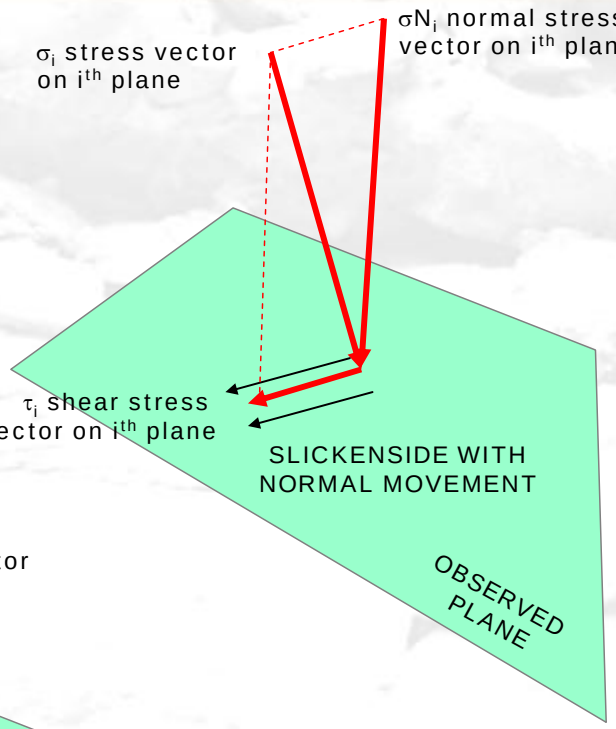
- **Normal stress** on the i[th] plane is the following:

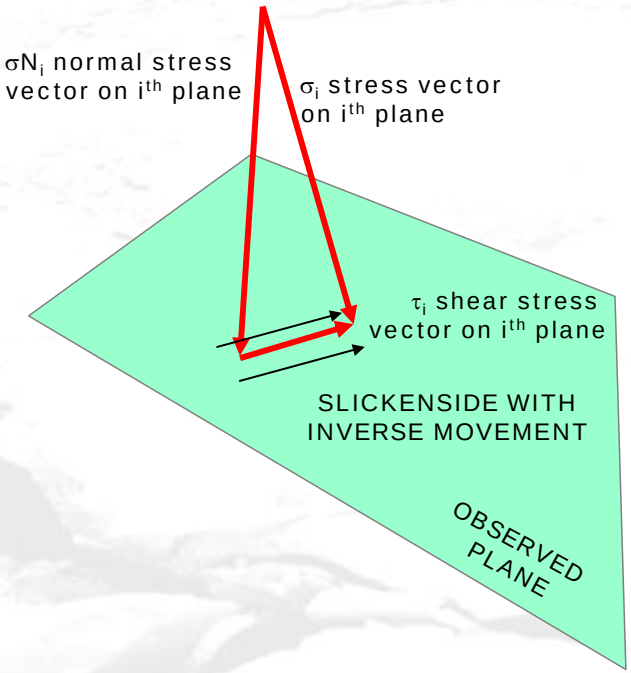
$$\sigma N_i = n_i \cdot \sigma_i = \begin{bmatrix} \sigma_i X \cdot n_1 \\ \sigma_i Y \cdot n_2 \\ \sigma_i Z \cdot n_3 \end{bmatrix}$$

$$\sigma_i = \begin{bmatrix} \sigma_i X \\ \sigma_i Y \\ \sigma_i Z \end{bmatrix}, n_i = \begin{bmatrix} NX \\ NY \\ NZ \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

- The **shear stress** on the i[th] plane is:

$$\tau_i = \sigma_i - \sigma N_i = \begin{bmatrix} \tau_i X \\ \tau_i Y \\ \tau_i Z \end{bmatrix} = \begin{bmatrix} \sigma_i X - \sigma N_i X \\ \sigma_i Y - \sigma N_i Y \\ \sigma_i Z - \sigma N_i Z \end{bmatrix}$$





STRESS VECTORS AND STRESS ESTIMATORS II.

υ VECTOR, λ, RUP AND ANG PARAMETERS

VECTOR υ

- Upsilon vector (υ) refers to the **divergence** between the **measured** (a priori, described by *DC* vector) and the **calculated** (a posteriori, τ vector) shear stress direction

$$\upsilon_i = DC_i \cdot \lambda - \tau_i = \begin{bmatrix} DCX \cdot \lambda - \tau_X \\ DCY \cdot \lambda - \tau_Y \\ DCZ \cdot \lambda - \tau_Z \end{bmatrix},$$

$$SV_i = \begin{bmatrix} DC_iX \\ DC_iY \\ DC_iZ \end{bmatrix}, \tau_i = \begin{bmatrix} \tau_X \\ \tau_Y \\ \tau_Z \end{bmatrix}$$

SHEAR STRESS VECTOR LENGTH λ

- Angelier (1990) assumed shear vector length to be √3/2; this is the 0th iteration step of Mostafa's (2005) methodology. λ is the scale factor to get the **real** length of the slip vector instead the unit vector length; in Angelier's (1990) methodology its value is constant √3/2, in Mostafa's (2005) methodology it's value changes iteration by iteration

ESTIMATOR ANG

- Estimator **ANG** shows the angular misfit between the measured (DC) and calculated (τ) shear stress vectors:

$$ANG = ang(s_i, \tau_i)$$

$$s_i = \begin{bmatrix} DCX \\ DCY \\ DCZ \end{bmatrix}, \tau_i = \begin{bmatrix} \tau_i X \\ \tau_i Y \\ \tau_i Z \end{bmatrix}$$

ESTIMATOR RUP

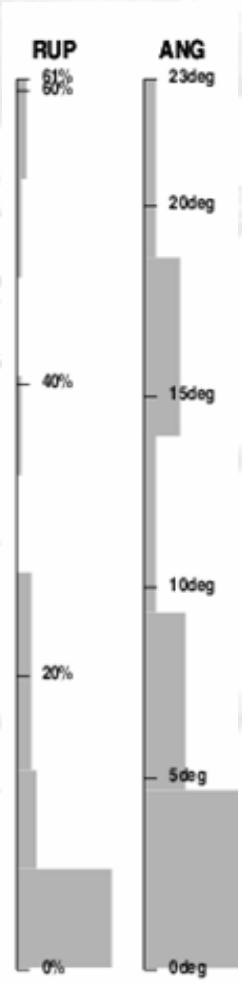
- Estimator **RUP** (**r**elative **up**silon) shows the calculated τ shear vector length relative to assumed λ = √3/2

$$RUP_i = \frac{\sqrt{\upsilon^2}}{\lambda} \cdot 100\% =$$

$$\frac{\sqrt{((\lambda^2 + \tau_i X^2 + \tau_i Y^2 + \tau_i Z^2) - 2 \cdot \lambda \cdot ANG)^2}}{\lambda} \cdot 100\%,$$

$$\upsilon^2 = \lambda^2 + \tau^2 - 2 \cdot \lambda \cdot ANG$$

RUP and ANG diagram for TYM data set (Angelier 1990) ➡



- Because absolute stress magnitudes are **unknown**, Mohr circle parameters were **re-scaled to unit stress**:
  - $\sigma_1 = 1.0$,
  - $\sigma_2 = \varphi$, and
  - $\sigma_3 = 0.0$

- Calculate the **stress vector magnitude:**

$$|\sigma| = n_1\sigma_X + n_2\sigma_Y + n_3\sigma_Z,$$

$$\sigma = n_i \cdot T = \begin{bmatrix} \sigma_X \\ \sigma_Y \\ \sigma_Z \end{bmatrix} = \begin{bmatrix} n_1\sigma_{11} + n_2\sigma_{12} + n_3\sigma_{13} \\ n_1\sigma_{12} + n_2\sigma_{22} + n_3\sigma_{23} \\ n_1\sigma_{13} + n_2\sigma_{23} + n_3\sigma_{33} \end{bmatrix}$$

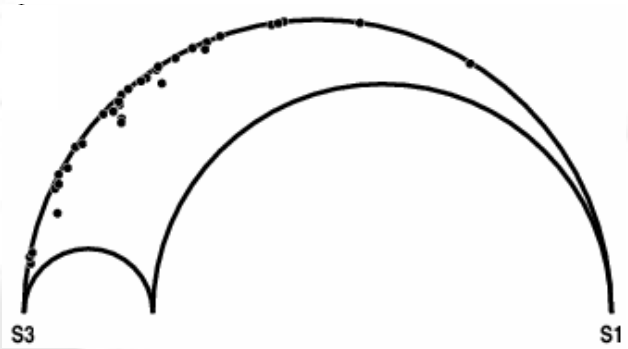
- **Normal stress value** will be computed on the following way:

$$\sigma N_i = \frac{\sqrt{\sigma N_i X^2 + \sigma N_i Y^2 + \sigma N_i Z^2} - \sigma_3}{\sigma_1 - \sigma_3}$$

- **Shear stress** on the i^{th} plane is as follows:

$$\tau_i = \frac{\sqrt{\tau_i X^2 + \tau_i Y^2 + \tau_i Z^2}}{\sigma_1 - \sigma_3}$$

- Little black thick on the colored stress bar refers to stress regime computed after Delvaux et al's (1999) reduced stress tensor



Mohr circle for TYM data set (Angelier 1990)

R' = 2.75 – 3.00: radial compressive

R' = 2.25 – 2.75: pure compressive

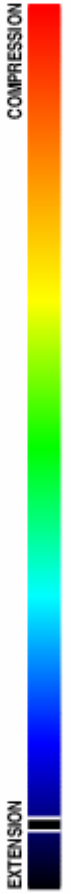
R' = 1.75 – 2.25: transpressive

R' = 1.25 – 1.75: pure strike slip

R' = 0.75 – 1.25: transtensive

R' = 0.25 – 0.75: pure extensive

R' = 0.00 – 0.25: radial extensive



CONTOURING METHODOLOGY

TRIANGULAR GRID

- Generate a **Kaalsbeek counting net** (331 points), and convert in to Schmidt / Wulff net
- Identify **individual triangles** of the counting net (660 triangles)
- Decide which **input values** to count (dip, dip direction, strike, bearing, plunge), **count** data number in each triangle, and store them as a triangle center – count data set

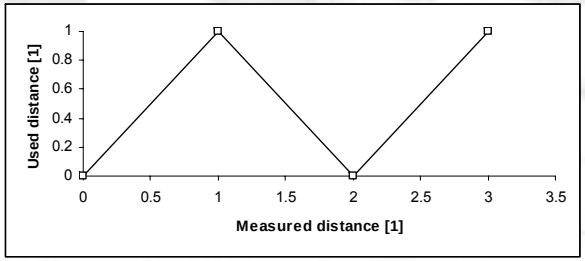
RECTANGULAR GRID

- Generate a **rectangular grid** with the same number of data points in the unit circle as the Kaalsbeek triangle net has – use approximately √(331 / 0.78) cells
- **Re-scale** triangular grid's count data to the rectangular grid, calculating the count of each rectangular grid point C_R using each triangular counts C_T, where R_X, R_Y, T_X and T_Y are the rectangular and triangular grid point coordinates, respectively, and d is the distance of C_R and C_T

$$C_R = \frac{C_T}{d},$$

$$d = \sqrt{(R_X - T_X)^2 + (R_Y - T_Y)^2}$$

- If a data set is close to the unit circle of the stereonet, the contour lines should appear on the opposite side of the stereonet as well because the projection method. To be able to visualize these contour lines, **re-calculate the measured distance**



- To avoid division by zero when C_R and C_T are close to each other, **define a minimum value for d**
- Store as grid point – count data set
- Define **contouring values V** as a percentage of the maximum value of the rectangular grid

BINARY GRID

- Generate **binary rectangular grid** using the original rectangular grid: cell value is 1 if the value in the cell is higher that the value of the contour line, otherwise keep it zero
- Use **marching squares** methodology to generate line segments, and re-calculate their relative position taking the contouring percentage into agreement

LINE SEGMENT PROCESSING

- If a line is out of the unit circle, **ignore** it
- If a line crosses the unit circle, **clip** it
- If a line is not closed and does not cross the unit circle, linearly **extrapolate** first and/or last vertex coordinate to be at the unit circle
- **Eliminate** short lines (e.g. less than 5 vertex)

LINE PROCESSING

- If a line terminates at the unit circle, it has to be **closed** either in the first → last or last → first vertex direction (nothing to do with the closed lines)
- Generate **two curves**, one is closed into the first → last vertex direction, the other one in the last → first vertex direction
- Each contour line has a value V (see above): **if any of the two lines** contains any grid point count value equal or higher than V, than use that line
- If **two contour lines** has any grid point equal or higher than V, than choose the one with the **smaller** area
- In **none** of the two contour lines has any grid point equal or higher than V, that's a bug
- Generate **2nd order Bezier curves** from the line segments, and plot them

7 REFERENCES

- **Angelier, J. 1979**: Determination of the mean principal stresses for a given fault population. – Tectonophysics 56, T17–T26
- **Angelier, J. 1984**: Tectonic analysis of fault slip data sets. – Journal of Geophysical Research 89, 5835–5848
- **Angelier, J. 1990**: Inversion of field data in fault tectonics to obtain the regionalstress III: A new rapid direct inversion method by analytical means. – Geophysical Journal International 103, 363–376
- **Fry, N. 1999**: Striated faults: visual appreciation of their constraint on possible palaeostress tensors – Journal of Structural Geology 21, 7–27
- **Michael, A. 1984**: Determination of stress form slip data: fault and folds, Journal of Geophysical Research 89,, 11517-11526
- **Mostafa, M. E. 2005**: Iterative direct inversion: An exact complementary solution for inverting fault-slip data to obtain palaeostresses. – Computers & Geosciences 31, 1059–1070
- **Sasvári, Á. & Baharev, A. 2014**: SG2PS (structural geology to post script converter) – A graphical solution for brittle structural data evaluation and paleostress calculation. – Computers & Geosciences 66, 81–93
- **Shimazaki, H. & Shinomoto, S. 2007:** A method for selecting the bin size of a time histogram. Neural Computation 19/6, 1503-1527
- **Spang, J. H. 1972**: Numerical method for dynamic analysis of calcite twin lamellae. – Geological Society of America Bulletin 83, 467-472
- **Turner, F. J. 1953**: Nature and dynamic interpretation of deformation lamellaein calcite of three marbles. – American Journal of Science 251, 276 – 298

8 COPYRIGHT