Trabajos de Geología, Universidad de Oviedo, 29 : 292-296 (2009)

Oviedo (Spain) 1-3 July 2008 International Meeting of Young Researchers in Structural Geology and Tectonics
ORSGET

Performance maps – a tool for examination of reliability of procedures for automatic separation of heterogeneous fault/slip and calcite twin data

Ł. Gągała*

Institute of Geological Sciences, Jagiellonian University. Ul. Oleandry 2a, 30-063 Kraków, Poland.

**e-mail: diopsyd@o2.pl*

Abstract: The concept of performance maps is introduced in order to investigate a reliability of routines for automatic separation of polyphase fault/slip and calcite twin data. They show a performance of a given procedure in relation to a difference between superposed stress tensors and proportion between sizes of homogeneous subsets. The reliability of two procedures is investigated with performance maps – sequential separation for fault/slip data and s_1 clustering for calcite twins. The results suggest a high reliability of the sequential separation and poor performance of σ_1 clustering.

Keywords: fault/slip data, calcite twins, data separation, stress inversion.

The quantitative dynamic interpretations in weakly deformed strata are usually performed by stress tensor inversion from fault/slip data (Etchecopar *et al.*, 1981; Angelier, 1984) and less frequently from deformation twins in calcite crystals (Etchecopar, 1984 *fide* Tourneret and Laurent, 1990). The most important complication encountered during processing natural data sets is their polyphase nature. Since they are usually a product of several superimposed stress states, a computation of single stress tensor for the whole data set is erroneous. In case of striated faults this difficulty may be often overcome by an extraction of dynamically homogeneous subsets upon a basis of qualitative criteria; e.g. morphology and mineralization of slickenside surface, several generations of slickenside lineation, cross-cutting relationships between faults etc., but for calcite twins this method is not applicable, since these microstructures do not bear any signature of their chronological order or reactivation and the only method for identification of homogeneous subsets is an automatic separation upon a basis of dynamic criteria (Etchecopar, 1984 *fide* Tourneret and Laurent, 1990; Rocher *et al.*, 2004). This approach is also applicable for complex, morphologically uniform populations of striated faults, when qualitative (or "field") criteria cannot be reliably applied (Lacombe *et al.*, 1990 and 1992; Nemcok and Lisle, 1995; Nemcok *et al.*, 1999; Yamaji, 2000; Rocher *et al.*, 2004; Yamaji *et al.*, 2006; Otsubo *et al.*, 2006).

Several routines were proposed for fault/slip and calcite twin data separation. All of them take advantage of dynamic inconsistencies between homogenous subgroups forming a polyphase data set. Two main classes of separation procedures may be distinguished: separation contemporaneous with inversion and separation prior to inversion. The first philosophy is applied for calcite twins while the latter both for calcite twins and striated faults. All of the hitherto proposed routines were validated with natural and several of them also with numerical polyphase data sets. The proofs of validity are usually single case studies, where only one data set is analyzed, reflecting a unique configuration of stress states. However, the experience with natural and numerical examples shows that results of data separation are highly dependent on a structure of data set (relative sizes of homogeneous subgroups) and similarity of superposed stress states, thus the procedure well performing for some configurations of these factors is likely to fail for others. Therefore, it is postulated, that a proper evaluation of a performance of any method for data separation should be based on its systematic application to a possibly wide range of data sets' structures and configurations of superposed stress tensors.

In this contribution, a performance of two separation routines developed by the author is systematically investigated over a full range of configurations of superposed stress states and structures of data sets. These are: a method of separation contemporaneous with inversion for striated faults, based on an algorithm of Etchecopar (1984 *fide* Tourneret and Laurent, 1990) developed originally for calcite twins (proposed name: sequential separation) and a method of separation prior to inversion for calcite twins, based on clustering of σ_1 calculated separately for each twin (proposed name: σ_1 clustering). The readers are referred to an original paper (Gągała, *in press*) for details of these algorithms.

Methods

The data generation, separation and stress inversion was performed iteratively for various configurations of stress tensors and proportions between homogeneous data subsets. Only two-phase data sets not contaminated by chaotic (incompatible) admixture were investigated. This is a simplification with respect to natural data, but allows an evaluation of the highest possible performance of given procedure. The process may be summarized in the following main steps:

1. Generation of two random complete stress tensors such as σ_1<100 MPa and σ_3>-10 MPa. The absolute values of principal stresses are recalculated to a tensor shape ratio, according to a convention:

$$\Phi = \frac{\sigma_2 - \sigma_3}{\sigma_1 - \sigma_3} \qquad (1).$$

The difference between stress tensors is expressed according to Lisle and Orife (2002) as an octahedral shear stress D. The range of this parameter is (0, 2) for identical and negative tensors respectively. It must be noted that D is not a linear function of input parameters (rotation angles and principal stresses for both tensors). A uniform distribution of these parameters results in a non-uniform distribution of computed stress differences D, where values close to 0 and 2 are poorly represented.

2. Generation of random potential faults and twin planes (200 and 300 respectively) and solving a direct problem by a successive application of two stress tensors. As a result the following arrays are produced:

a) Arrays with normalized plane normals and slip vectors for numerical fault/slip data.

b) Arrays with normalized plane normals and slip vectors separately for twinned planes and potential, not twinned planes for numerical calcite twin data.

The ratio between numbers of numerical faults or twinned planes compatible with each tensor is expressed as:

$$N = \frac{n_1}{n_2} \qquad (2),$$

where n_1 and n_2 are absolute sizes of homogeneous subgroups, such as n_1 corresponds to the more numerous subgroup regardless its compatibility with the first or the second stress tensor. The N parameter indirectly reflects a stress difference D.

3. Separation of two-phase data sets with the investigated separation algorithm and stress inversion from extracted, potentially homogeneous subgroups.

4. Verification of solutions obtained in point 3 by a comparison of computed and imposed stress tensors according to arbitrary criteria of correctness. For directions of principal stresses a 10° deviation is allowed, while for shape factor Φ a tolerance of 0.1 is admitted.

5. Construction of performance maps. Points 2-4 are repeated 10 times for a single pair of stress tensors. Obtained N values (10) are averaged and a number of successful separations is stored as a V parameter (ranging from 0 to 100% in increments of 10%). Thus, after several iterations of the whole procedure an array is produced, containing triplets (D, N, V). The grid of uniformly spaced D and N is created; V values are interpolated at grid nodes and averaged over grid cells of sizes set arbitrarily (0.3×1.5 for D and N respectively). Obtained results are graphically summarized in performance maps (Figs. 1a and 2a), produced from triplets (D, N, V). They show a percentage of successful trials V in relation to stress difference D and average proportion between homogeneous subsets N.

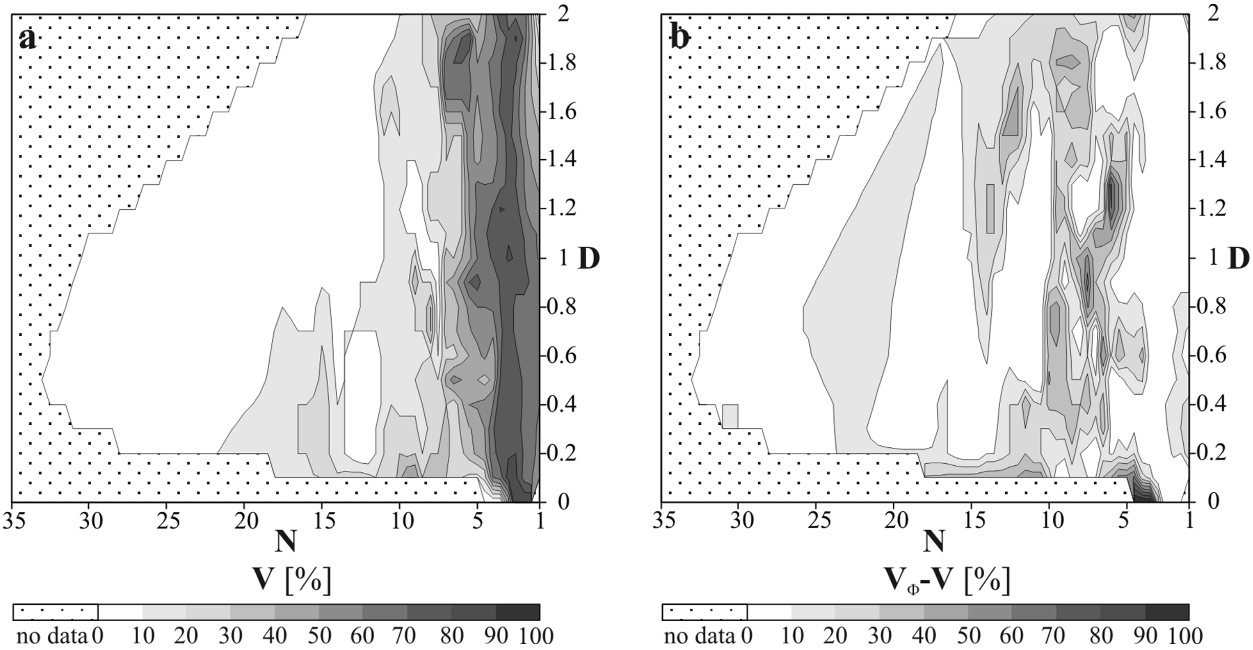

Figure 1. (a) Performance map of sequential separation. No sensitivity to D is detected. Reliability is satisfactory for N<5, (b) map of differences V_Φ-V. Inaccurate calculations of Φ are an important source of errors in stress inversion.

Normally, the accuracy of solutions is evaluated upon a basis of rotation angles and shape factors of computed stress tensors. However, the verification (point 4) may be also done exclusively by a comparison of imposed and computed rotation angles without taking into account the shape factor Φ. The performance maps constructed from numerical experiments carried out with these more tolerant criteria of correctness, as a rule, show higher reliability (further referred to as V_Φ) than those produced for the full set of crite-

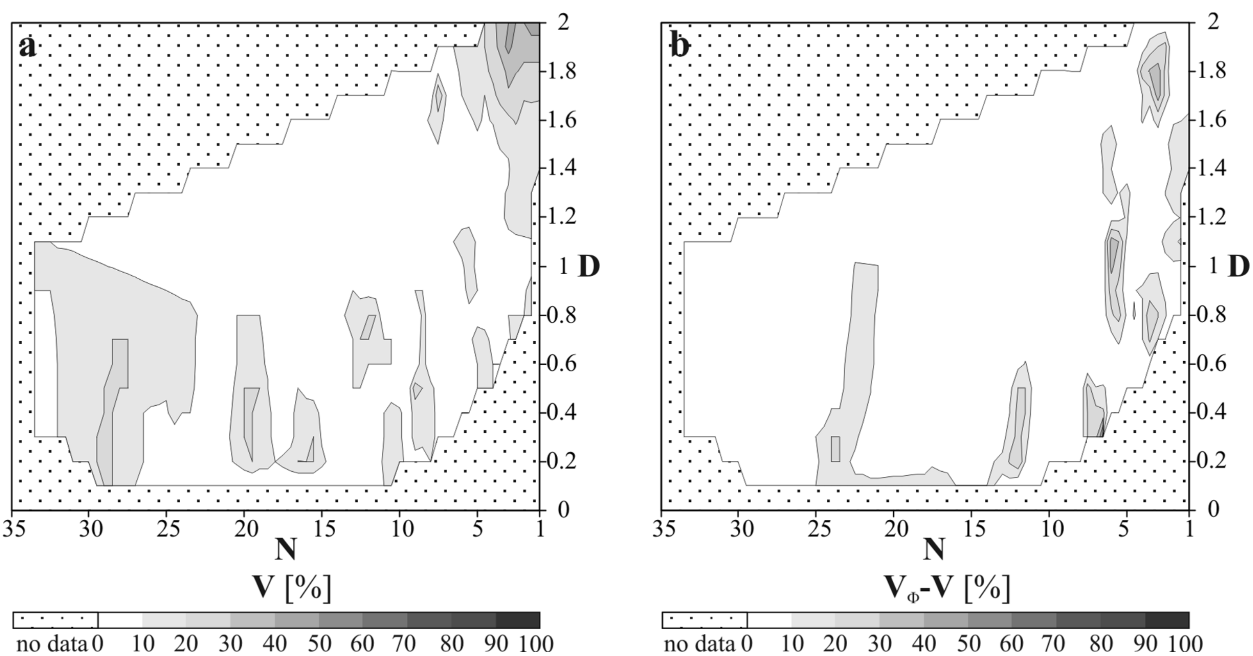

Figure 2. (a) Performance map of s_1 clustering. General poor performance and sensitivity both to D and N is inferred, (b) map of differences V_Φ-V. The Φ factor appears as an unimportant source of errors in stress inversion.

ria (V). If these two types of performance maps are subtracted one from another, the map of differences is created from triplets (D, N, V_Φ-V), showing an extent to which an inaccurate evaluation of Φ contributes to classification of the whole solution as unacceptable (Figs. 1b and 2b).

In a course of numerical experiments more than 400 D values were examined. Each procedure was tested twice: with complete and reduced (without Φ) criteria of correctness. Thus, more than 8000 trial separations were performed for each separation algorithm.

Results

The performance maps for sequential separation for fault/slip data and σ_1 clustering for calcite twins are shown in figure 1a and 2a respectively. It is clear that in both cases statistically satisfactory performance is obtained for restricted intervals of D and N.

Sequential separation

The performance map for sequential separation (Fig. 1a) documents a low sensitivity of this procedure to a stress difference D. The main factor influencing its reliability is a relation between numbers of data in particular homogeneous subsets. For N in the range (0, 5) it may be considered satisfactory, since areas with V>60% predominate. The 100% reliability was documented for several points (configurations of D and N) in the discussed area of performance map. For N>5 the performance of the discussed procedure is not acceptable, despite of several clusters, where it is evidently elevated with respect to the background (i.e. for 1.6<D<1.9 and 6<N<8).

The map of differences (V_Φ-V) documents a high influence of inaccurate Φ calculations upon correctness of the whole solutions. Areas with differences higher than 40% occupy a significant part of the plot for N between 5 and 10.

σ_1 clustering

The performance map for s_1 clustering (Fig. 2a) documents a general low reliability of this procedure and its sensitivity both to a stress difference D and proportion between homogeneous subsets N. The highest level of average reliability (40-50%) is limited to extreme values of D>1.9 and N<3. In this range of D and N infrequent experiments with 90% reliability were noted, but they may be considered as exceptions only.

The map of differences (V_Φ-V) shows rather low influence of inaccurate assignations of Φ upon correctness of complete tensor solutions. The V_Φ higher than V of 10 up to 40% is noted for small areas of difference map for N between 1 and 7. Clusters of difference in performance between 10 and 20% exist also for N between 10 and 25.

Discussion

The sequential separation of polyphase fault/slip data may be considered as a well performing routine. The limitation of applicability for N<5 is not a significant problem, since it is rarely possible to collect more than 50 fault/slip measurements from a single site. In the extreme situation of N=5, the less numerous homogeneous population would comprise between 8 or 9 measurements only, which is not enough for a reliable paleostress inversion anyway (Orife and Lisle, 2006). On the other hand, low sensitivity of this procedure to a stress difference D facilitates separation of subsets originated due to similar stress states, which is a strong advantage. The danger of inaccurate computation is inherent in the inversion method for fault/slip data. Since it is based exclusively upon directional data (real and computed slip directions) the shape factor of stress tensor may be weakly constrained if orientation of faults is preferential or the data set is contaminated by incompatible outliers of the less numerous homogeneous population. Thus, for a better control of this parameter, another separation method (Nemcok *et al.*, 1999; Yamayi *et al.*, 2006) should be used along with the sequential separation and preferably, the extracted homogeneous subsets should be checked against qualitative field evidence.

The σ_1 clustering did not prove its validity as it was postulated by author (Gągała, *in press*), thus this routine should be further elaborated or abandoned. The very small area of acceptable reliability rather excludes its application to natural samples. This limitation is even more important than for fault/slip data, since very often a number and dynamics of superposed paleostress states cannot be independently assessed in order to provide a control of obtained results. Additionally, on contrary to fault/slip data, the structure of data set cannot be verified qualitatively. The relatively low influence of incorrect assignations of Φ upon the whole solutions results from a structure of an inversion algorithm, which takes into account normalized minimum shear stress for twinning, along with directional data. Thus, the stress tensor parameters are better constrained than for striated faults. In the light of these considerations the σ_1 clustering may be applied at most as an additional check of the more reliable procedure i.e. original algorithm of Etchecopar (1984 *fide* Tourneret and Laurent, 1990).

Conclusions

1. The performance maps are a potential tool for evaluating applicability of procedure for separation of heterogeneous fault/slip and calcite twin data sets.

2. The investigated procedures for polyphase fault/slip calcite twin data show limited fields of applicability

3. The sequential separation may be considered as a reliable routine for fault/slip data separation for populations, where proportion between sizes of homogeneous data sets is not higher than 5. It is also not sensitive to a stress difference, which is a strong advantage. Stress inversion from automatically separated homogeneous data sets may return an inaccurate value of Φ.

4. The σ_1 clustering for calcite twins displays a very limited area of applicability and a strong sensitivity both to stress difference and proportion between relative sizes of homogeneous data sets. Stress inversion from automatically separated homogeneous data sets usually returns correct value of Φ.

5. The limited fields of applicability of routines for data separation encourage to simultaneous application of various algorithms to a single natural data set in order to provide mutual control of obtained solutions. Preferably they should be independently verified by qualitative field data.

References

ANGELIER, J. (1984): Tectonic analysis of fault slip data sets. *J. Geophys. Res.*, 89, B7: 5835-5848.

ETCHECOPAR, A., VASSEUR, G. and DAIGNIÉRES, M. (1981): An inverse problem in microtectonics for the determination of stress tensors from fault striaton analysis. *J. Struct. Geol.*, 3, 1: 51-65.

ETCHECOPAR, A. (1984). *Etude des états de contraintes en tectonique cassante et simulations de déformations plastiques (approche matématique).* Thèse d'Etat, USTL Montpellier, 269 pp.

GĄGAŁA, Ł. (in press): Reliability of selected procedures of stress inversion and data separation for inhomogeneous populations of calcite twins and striated faults – insights from numerical experiments. *Int. J. Earth Sci.*, doi:10.1007/s00531-007-0262-3.

LACOMBE, O., ANGELIER, J., LAURENT, P., BERGERAT, F. and TOURNERET, C. (1990): Joint analyses of calcite twins and fault slips as a key for deciphering polyphase tectonics: Burgundy as a case study. *Tectonophysics*, 182: 279-300.

LACOMBE, O., ANGELIER, J. and LAURENT, P. (1992): Determining paleostress orientations from faults and calcite twins: a case study near the Sainte-Victiorie Range (southern France). *Tectonophysics*, 201: 141-156.

LISLE, R. J. and ORIFE, T. (2002): STRESSTAT: a Basic program for numerical evaluation of multiple stress inversion results. *Comput. Geosci.*, 28: 1037-1040.

NEMCOK, M. and LISLE, R. J. (1995): A stress inversion procedure for polyphase fault/slip data sets. *J. Struct. Geol.*, 17, 10: 1445-1453.

NEMCOK, M., KOVÁC, D. and LISLE, R. J. (1999): A stress inversion procedure for polyphase calcite twin and fault/slip data sets. *J. Struct. Geol.*, 21: 597-611.

ORIFE, T. and LISLE, R. J. (2006): Assessing the statistical significance of palaeostress estimates: simulations using random fault-slips. *J. Struct. Geol.*, 28, 6: 952-956.

OTSUBO, M., SATO, K. and YAMAJI, A. (2006): Computerized identification of stress tensors determined from heterogeneous fault-slip data by combining the multiple inverse method and k-means clustering. *J. Struct. Geol.*, 28, 6: 991-997.

ROCHER, M., CUSHING, M., LEMEILLE, F., LOZAC'H, Y. and ANGELIER, J. (2004): Intraplate paleostress reconstructed with calcite twinning and faulting: improved method and application to the eastern Paris Basin (Lorraine, France). *Tectonophysics*, 387: 1-21.

TOURNERET, C. and LAURENT, P. (1990): Paleo-stress orientations from calcite twins in the North Pyrenean foreland, determined by the Etchecopar inverse method. *Tectonophysics*, 180: 287-302

YAMAJI, A. (2000): The multiply inverse method: a new technique to separate stresses from heterogeneous fault-slip data. *J. Struct. Geol.*, 22: 441-452.

YAMAJI, A., OTSUBO, M. and SATO, K. (2006): Paleostress analysis using the Hough transform for separating stresses from heterogeneous fault-slip data. *J. Struct. Geol.*, 28: 980-990.