The utility of statistical analysis in structural geology

Nicolas M. Roberts[1*], Basil Tikoff[1], Joshua R. Davis[2], Tor Stetson-Lee[1]

[1]Department of Geoscience, University of Wisconsin—Madison. 1215 West Dayton St., Madison, WI 53706*

[2]Department of Mathematics and Statistics, Department of Computer Science, Carleton College, Northfield

* nmroberts@wisc.edu

* *(608) 262-4678*

Key words: structural geology; directional statistics; orientation statistics; western Idaho shear zone; inference; regressions

Abstract

Recent advances in statistical methods for structural geology make it possible to treat nearly all types of structural geology field data. These methods provide a way to objectively test hypotheses and to quantify uncertainty, and their adoption into standard practice is important for future quantitative analysis in structural geology. We outline an approach for structural geologists seeking to incorporate statistics into their workflow using examples of statistical analyses from two locations within the western Idaho shear zone. In the West Mountain location, we test the published interpretation that there is a bend in the shear zone at the kilometer scale. Directional statistics on foliations corroborate this interpretation, while orientation statistics on foliation-lineation pairs do not. This discrepancy leads us to reconsider an assumption made in the earlier work. In the Orofino location, we present results from a full statistical analysis of foliation-lineation pairs, including data visualization, regressions, and inference. These results agree with

thermochronological evidence that suggests that the Orofino area comprises two distinct, subparallel shear zones. The R programming language scripts that were used for both statistical analyses can be downloaded to reproduce the statistical analyses of this paper.

1. Introduction

Structural geologists routinely work with datasets that are logistically limited to small sample size and/or spatial extent. When working with such data, an important—but under-appreciated—task should be to determine what can reasonably be interpreted about the geologic system in question. This determination depends on the uncertainty that arises because the dataset is an incomplete representation of the larger system. The field of statistics is fundamentally concerned with this data-to-system uncertainty, and statistical methods have important utility for any empirical research. As structural geologists, we can use statistics to better identify trends, understand mean(s) and dispersion in datasets, test hypotheses, evaluate implicit assumptions, and communicate the confidence of our interpretations to peers.

In most publications, structural geologists make interpretations using quantitative data (e.g. fabric measurements) and qualitative estimates of uncertainty. The lack of statistical treatment of structural geology data is in part a historical issue: there is not a strong tradition of training structural geologists in statistics. As a discipline born out of field studies and geologic mapping, early structural geology methods—including quantitative ones—developed without a statistical framework. Even with the eventual development of such a framework for *directional* (rays and lines) data types such as paleomagnetic poles, lineations, poles to foliation, paleocurrents, and fault striations (e.g. Davis and Sampson, 1986; Ducharme et al., 1985; Fisher et al., 1989; Jupp and Mardia, 1989; Merrett and Allmendinger, 1990; Yonkee and Weil, 2015), the statistically savvy structural geologist is still unusual. Although contouring directional data has become commonplace as a result of computer programs such as Stereonet (version 10.0.0; Allmendinger, 2017) and Orient (version 3.6.3; Vollmer, 2017), structural geologists do not generally make use of directional statistics to report statistical descriptors (mean, dispersion) or perform hypothesis tests in a statistically rigorous fashion.

Another reason structural geologists do not generally employ statistics is that many geologic data are not rays or lines, and thus cannot be treated with directional statistics. Until recently, there was no unified framework for the statistical treatment of *orientation* (line-within-plane) data like foliation-lineation pairs, fault planes with slickenlines, axial planes with hinges, or focal mechanisms. Davis and Titus (2017) have developed the mathematical background and theory of orientation statistics in a manner accessible by structural geologists. Moreover, they developed a free R programming language library for both direction and orientation statistics, called *geologyGeometry* (download the latest version at: http://www.joshuadavis.us/software/). Tools in the library include advanced plotting, regression algorithms, and parametric and non-parametric methods for inference, including hypothesis testing. The *geologyGeometry* library calls on other R libraries, including *Directional* (Tsagris and Athineou, 2016).

This contribution describes how to perform statistical analysis on structural geology data, and illustrates why incorporating statistics into a structural geology workflow is critical to the future of structural geology. Statistical analysis of two datasets from the western Idaho shear zone system, Idaho, USA are described in detail. The datasets were chosen because they address common questions in structural geology. A cursory geologic context is provided for each dataset (see Appendix 1 for additional details). The analysis of these two datasets reveals how thinking statistically leads to a more objective approach to interpretations and a quantified understanding of the uncertainty surrounding these interpretations. By demonstrating both its methodology and utility, we hope to motivate the adoption of statistics into the standard structural geology workflow.

All statistical analyses were done with the *geologyGeometry* R library, and the full analyses are shared in the appendices. Readers are encouraged to download the static version of the *geologyGeometry* library, which includes the datasets and statistical analyses from this contribution, at http://nicolasmroberts.github.io/scripts and run the scripts line-by-line. The scripts will output versions of all the data plots found in the figures, some of them interactive and rotatable.

2. The statistical approach

In the statistical analysis applied to each dataset in this paper, we utilize a new workflow motivated by both statistical protocol and geologic expertise. This workflow highlights two types of questions in statistics that are particularly relevant for structural geologists. First, are two datasets or subsets of a single dataset (e.g. two geographic domains) sampled from the same population? Second, are there real, systematic trends in the data, based on geographic position or any other variable?

Figure 1 presents a diagram of this workflow. First, the structural geologist visualizes the data in a variety of plots and maps. As a result of visualization, the geologist makes hypotheses and qualitative interpretations about the geologic system under study (lower path in Fig. 1), from which possible conceptual models and predictions are developed. Simultaneously, a statistical protocol is executed that is necessary for any dataset (upper path). Model predictions may be objectively tested by regressions (grey arrows), in which case the upper and lower path overlap. If there are no systematic spatial tendencies in the data, then the data can be statistically described by mean and dispersion, and inferences can be made about how well the population mean is known (the uncertainty of the population mean). The upper and lower paths interact again when model predictions are tested using statistical hypothesis testing. A statistical hypothesis test is formulated as a null hypothesis (e.g., “There is no difference in the mean of the populations from which dataset *A* and dataset *B* were sampled”) and an alternative hypothesis (e.g., “The means of the populations from which dataset *A* and dataset *B* were sampled do not have the same mean”). The null hypothesis is rejected or fails to be rejected based upon a credibility or confidence threshold, commonly 95%, or a p-value threshold, usually $p < 0.05$. A rejection of the null hypothesis leads a structural geologist to conclude that dataset *A* and dataset *B* are sampled from different populations. Importantly, the failure to reject the null hypothesis would *not* lead a structural geologist to conclude that dataset *A* and dataset *B* come from the same population—only that there is not strong evidence that they came from different populations.

Interpretations about the geologic system all pass through the statistical portion of the workflow, but statistics are only useful insofar as they complement geologic expertise to develop statistically constrained conclusions about the geologic system. The first example in this paper focuses on the “Mean, dispersion”, “Inference about the mean”, and “Statistical hypothesis testing” boxes of the statistical workflow (Fig. 1) to statistically test a published geologic interpretation. The second example illustrates a path through the workflow (as shown by thick lines in Fig. 1).

This statistical approach is compatible with a range of grouping parameters. We perform statistical tests on data grouped by geographic location, but these same methods can be performed on data grouped by relative age, composition, or any other parameter.

2.1 Uncertainty in data collection

An important source of uncertainty comes from how structural geologists sample data. Data are often not homogeneously distributed over the field area because of the availability of outcrop, and often there are structural controls on where outcrops occur. Rigorous treatment of this spatial sampling problem is difficult and outside the scope of this paper. We use datasets that were collected in areas of mostly contiguous outcrop.

Another source of uncertainty comes from the error of the devices and the measurement process. This topic is especially timely in the era of mobile devices (Allmendinger et al., 2017; Novakova and Pavlis, 2017). We assume that there is no systematic measurement error (bias), and we treat random measurement errors implicitly, together with other sources of variability in data. For example, suppose that we wish to characterize the average foliation in a field area, in which we have many foliation measurements, which vary by 20° or more. Perhaps 2° of this variation arises from the compass and the geologist user. Separating this source of variation from the much larger variation in the rocks complicates the analysis, probably with no appreciable effect on our confidence regions or conclusions. Multiple measurements per site can avoid this type of error, as the software of Allmendinger et al. (2017) does. That approach may particularly improve studies that would otherwise have very few data and is important for testing devices that may have systematic error associated with them, such as mobile devices. In this contribution, all data were collected with a traditional Brunton compass. Some sites have multiple measurements recorded, while only a single measurement was recorded at many sites.

3. Background: direction and orientation data

The mathematical description of a geologic structure's orientation depends on the type of geologic structure. A lineation can be described by two angles, a trend and plunge, or by a vector in Cartesian (east, north, up) coordinates. A foliation can also be described by two angles—a strike and a dip—and similarly can be described as a single Cartesian vector that defines the pole to the foliation. Foliations and lineations are both examples of *directional* data, meaning that a single line or ray is sufficient to uniquely describe the geometry. A wealth of statistical techniques have been developed for directional data (e.g. Mardia and Jupp, 2000), termed *directional statistics*.

In contrast, a foliation-lineation pair is defined by a line oriented within a plane. At a minimum, three angles are required to describe a unique foliation-lineation: a strike, a dip, and a rake. Foliation-lineation pairs are an example of *orientation* data and are treatable by *orientation statistics* (Downs, 1972; Davis and Titus, 2017). For statistical treatment, orientation data can be represented by a 3 by 3 rotation matrix, with the first row comprising the Cartesian vector of the pole to foliation, the second row comprising the vector of the lineation, and the third row comprising the vector that is orthogonal to the first two rows (Davis and Titus, 2017).

4. Application 1: The western Idaho shear zone near West Mountain, ID

The western Idaho shear zone forms a steep and abrupt north-south boundary between accreted terranes and the cratonic edge of the North American Cordillera (Armstrong et al., 1977; Fleck and Criss, 1985; Manduca et al., 1992; Tikoff et al., 2001; Fleck and Criss, 2004; Braudy et al., 2017). The ~5 kilometer wide shear zone is characterized by highly deformed orthogneisses. In the West Mountain area, the western Idaho shear zone is dextral, but sub-vertical lineations suggest transpression (Giorgis et al., 2008; Giorgis et al., 2016). The shear zone system bends at the 100-km scale to follow the cratonic boundary as defined by the $^{87}Sr/^{86}Sr$ isopleth (Figure 2A inset). Miocene extension resulted in a (~10°) east-west tilt, so that originally vertical foliation now dips at approximately 80° (Tikoff et al., 2001).

A recent structural study suggests that a subtle bend in shear zone orientation can also be detected at the kilometer scale near West Mountain, ID. Braudy et al. (2017) collected a dataset of field fabrics including both foliation-only measurements as well as foliation-lineation pairs (Fig. 2B and 2C). They plot both types of fabric data in equal area nets and interpret a ~20° rotation between foliation strike in the North and South of the field area. In this section, we demonstrate a statistical analysis that is able to provide a more objective test of the interpretation of Braudy et al. (2017). In addition, we test the assumption that the foliations from the foliation-only dataset and the foliation-lineation dataset are the same. Foliation-only and foliation-lineation pairs are treated separately because they are different data types.

4.1 Directional statistics on foliation-only data

Foliation-only data comprise 148 field fabric measurements. Braudy et al. (2017) divide these data into three geographic domains: northern (n = 56), central (n = 23), and southern (n = 69) (Fig. 2B). For the sake of comparison, these domain divisions are used in this statistical analysis.

For each domain, the data are used to make an inference about the population mean. This calculation is done by bootstrapping (Efron and Tibshirani, 1994) and by applying three two-sample tests for comparison. Bootstrapping is repeated sampling with replacement. The implementation of the bootstrapping routine in the *geologyGeometry* R library is straightforward to use and automatically computes confidence regions (see Appendix 2 for a simplified description, or Davis and Titus (2017) for a full description). The result of bootstrapping is a cloud of means, whose center approximates the mean of the dataset and whose density at any given point is related to the likelihood of that point being the population mean. The 95% confidence ellipse of each domain is calculated using the Mahalonobis distance (Mahalanobis, 1936).

To determine whether the foliations in each domain come from different populations, a series of statistical tests are devised. The null hypotheses are that the population mean of one domain (e.g. northern) is the population mean of another domain (e.g. southern). The null hypothesis is rejected if the 95% confidence regions of the two domains in question do not overlap.

Results of bootstrapping and 95% confidence region calculations are summarized in Figure 3A. The 95% confidence ellipses of the southern and central domains overlap, but neither of these domains overlap with the northern region. We fail to reject the null hypothesis that the southern and central domains are sampled from the same population. For both comparisons with the northern domain, we reject the null hypothesis of a single population at 95% confidence.

In addition to bootstrapping, three types of two-sample tests were applied to each pair of domains. See Appendix 2 for a brief description of each test. A Wellner test (Wellner, 1979) yields a $p < 0.0001$ based on 10,000 permutations for the northern and southern domains as well as for the northern and central domains. The p-value for the southern and central comparison is 0.427. Two variations of Watson inference tests were performed on the data (Mardia and Jupp, 2000), one that assumes tightly concentrated data and the other that assumes large sample size. Respectively, these tests yield p-values of 0.000001 and 0 (northern and southern domains), 0.00002 and 0.000005 (northern and central domains), and 0.233 and 0.131 (central and southern domains). These p-values agree well with the bootstrapping results, although some caution is advised, since these tests make assumptions about how the data are distributed. Taken together, these tests provide strong evidence against the null hypotheses that the northern domain and the central/southern domains are sampled from the same population.

Given the statistically significant difference between the northern domain and the other two domains, we calculate the rotational difference between the northern and southern domains. The axis and magnitude of rotation between the northern and southern domains is determined by computing the minimum rotations between 10,000 pairs of northern and southern bootstrap means. Results from this analysis are summarized in Figure 3A. The mean rotation is 12.20° ± 3.8° (2σ) with a mean rotation axis that trends 164.3° and plunges 68.8° (see Figure 3 for the 95% confidence ellipse). These results contrast with the interpretation of Braudy et al. (2017), who suggest a ~20° rotation and implicitly assume a vertical axis rotation.

4.2 Orientation statistics on foliation-lineation data

Foliation-lineation data comprise 129 field fabric measurements. The data are analyzed with the same geographic domains described above: northern (n = 16), central (n = 34), and southern (n = 79).

For each domain, both a bootstrapping method and a Markov chain Monte Carlo (MCMC) simulation (Davis and Titus, 2017) produce clouds of possible means from which a confidence region (for bootstrapping) or a credible region (for MCMC) can be computed. See Appendix 2 for a brief description. In general, for small sample sizes (n < 30), MCMC returns credible regions with accurate size, but the regions tend to be unrealistically isotropic. By contrast, bootstrapping returns more realistic anisotropic confidence regions, but the size of the region is consistently underestimated. Because of these complementary strengths and weaknesses, it is helpful to use both approaches.

The null hypotheses for foliation-lineation pairs are identical to those for foliations described previously. If the bootstrap/MCMC confidence/credible regions do not overlap, then the null hypothesis can be rejected at 95% confidence/credibility.

Figure 3B shows the results of both bootstrapping and MCMC. The null hypothesis that the northern and central domains are sampled from populations with the same mean cannot be rejected using MCMC, but can be rejected using bootstrapping at 95% confidence. The same is true for the null hypothesis with respect to the northern and southern domains. The null hypothesis that the southern and central domains are sampled from populations with the same mean cannot be rejected at 95% confidence/credibility. Because MCMC credible regions tend to have more accurate coverage rates than confidence regions produced from bootstrapping (see the numerical experiments of Davis and Titus, 2017) for small sample sizes, these analyses do not provide strong evidence that differences among the three domains are statistically significant. There is, however, weak evidence that the null hypothesis can be rejected for the northern domain with respect to the other two domains.

4.3 Comparing foliation-only and foliation-lineation data

The statistical analysis of foliation-only and foliation-lineation fabric data from Braudy et al. (2017) leads to two different interpretations. From the foliation-only data, a $12.55° \pm 3.30°$ (2σ) rotation between the southern/central and northern domains is inferred. From the foliation-lineation data, no such difference can be inferred with statistical significance. This discrepancy motivates a statistical comparison of these two datasets.

In a final comparison, only the foliations are used from the foliation-lineation data, so that directional statistics can be applied to both datasets. The null hypothesis for each domain is that the foliations from the foliation-lineation dataset are sampled from the same population as those from the foliation-only dataset. A comparison of the bootstrapped mean cloud for each domain (Fig. 3C) shows that the null hypothesis can be rejected with 95% confidence for the central and southern domains, but is not clearly rejected for the northern domain. This result is unexpected because foliation-only data and foliation-lineation data were collected in the same field area (similar extent and spacing) and were assumed to be sampled from the same population of fabric orientation (Fig. 2A).

4.4 Summary

The statistical analysis of foliation-only and foliation-lineation data in the West Mountain area of the western Idaho shear zone allows for interpretations with quantitative evaluation of uncertainty. In addition, statistical comparison between foliation-only and foliation-lineation data reveals that a basic assumption about the two datasets—that the foliation and foliation-lineation datasets are being sampled from the same population—may not be valid.

The interpretation that Braudy et al. (2017) make with respect to foliation-only differences between the southern/central and northern domains is reasonable. Our statistical analysis agrees that a rotation has occurred, with high confidence. However, it is not clear that the rotation axis was vertical, as Braudy et al. (2017) implicitly assume, especially because Miocene tilting post-dated the formation of any bend in the western Idaho shear zone. Our analysis explores the alternative assumption that the rotation was the smallest possible. Under this assumption, the confidence region for the rotation axis plunges steeply to the south and does not contain the vertical axis. The magnitude of rotation is smaller than was found by Braudy et al. (2017). In this way, statistics helps us to explicitly identify assumptions and investigate how those assumptions affect geologic interpretations.

In contrast to the foliation-only data, statistical analysis of foliation-lineation pairs does not corroborate the counterclockwise rotation of fabric from south to north proposed by Braudy et al. (2017). The difference in fabric orientation among the domains is not statistically significant.

Braudy et al. (2017) do not interpret foliation-lineation pairs independently of the foliation-only data. By combining the datasets, they assume that foliations from the two datasets were sampled from the same population within each domain. Three statistical comparisons of the foliation data from the two datasets within each domain reject this assumption with 95% confidence for the central and southern domains. All three domains of the foliation-lineation foliations plot in the gap between the northern and southern domains of the foliation-only foliations (Fig. 3C). There are several possible explanations for this discrepancy which motivate future work. One possibility is that SL-tectonites may be differently oriented than S-tectonites because of strain partitioning within the western Idaho shear zone, so that two distinct populations of fabrics are inter-layered over the same field extent. Another possibility is that rocks that have only foliations also have a more poorly developed fabric than rocks with obvious lineations, which may account for the larger spread of data. Whatever the case, new scientific questions arise from the statistical analysis that would not have been asked in the absence of the application of statistical methods.

5. Application 2: Ahsahka and Woodrat Mountain shear zones near Orofino, ID

The Ahsahka shear zone (Figure 4) is in structural continuity with the western Idaho shear zone, about 200 km north of the West Mountain area (Giorgis et al., 2017; Schmidt et al., 2017). The Ahsahka shear zone occurs within a 90° bend of the western Idaho shear zone system (Figure

4 inset) (e.g., Lewis et al., 2014). The current interpretation is that there is an older, parallel Woodrat Mountain shear zone in cryptic contact with the northeast boundary of the Ahsahka shear zone (Lewis et al., 2014, Schmidt et al., 2017). They may also be younger shear zones present regionally (McClelland and Oldow, 2004, 2007; Lund et al., 2007). For the purposes of this statistical analysis, we group the deformed rocks north of the mapped boundary of the Ahsahka shear zone (Fig. 4) as Woodrat Mountain shear zone.

A recent dataset collected by Stetson-Lee (2015) comprises foliation-lineation measurements from areas on either side of the cryptic boundary between what is currently mapped as the Ahsahka and Woodrat Mountain shear zones near Orofino, Idaho (Fig. 4). There is a generally NW-striking foliation throughout the field area. Cooling $^{40}Ar/^{39}Ar$ ages on hornblende, biotite, and muscovite suggest that rocks on either side of the inferred boundary between the Ahsahka and Woodrat Mountain shear zones have a protracted thermal history, and record at least two distinct events (Davidson, 1990).

The goal of this statistical analysis is to assess whether structural data support the current interpretation of two distinct shear zones. If geographic domains have statistically significant orientation differences, are these differences consistent with the current inferred boundary? This statistical analysis illustrates the proposed structural geology workflow (Fig. 1), with a particular emphasis on the "statistical protocol" path. First, the data are visualized through a variety of plots. Second, the data are tested for geographic trends using regressions, and are split into geographic domains as a result. Third, the domains are statistically described with mean and dispersion. Finally, the domains are compared using hypothesis testing. The statistical tests are motivated and informed by data from the literature (maps, cooling ages) and the conceptual model for shear zone boundary that arises from them.

5.1 Statistical analysis using orientation statistics

The Orofino dataset comprises 69 foliation-lineation pairs in three geographic areas: Domain 1 (n = 23), domain 2 (n = 14), and domain 3 (n = 32) (Fig. 5). The division of the data in this way is consistent with current interpretations of geologic boundaries and will be statistically tested.

Initial plots contain all the data, not yet divided into geographic domains (Fig. 5A). Equal-area projections and equal-volume plots with Kamb contours show that the foliation-lineation data have an approximately unimodal distribution in their orientation. However, coloring the data by geographic location reveals a non-random relationship between orientation and geography. For example, when the data are color-coded by northing, there are clear domains of yellows and reds (Fig. 5A). In map view, this geographic dependence is apparent; lineations in the south trend north-northwest, while in the north they trend east-northeast.

Before splitting the data into domains, it is critical to know whether this geographic dependency is systematic (i.e. can be described by a continuous function) or whether there are discrete differences of orientation in different geographic domains. A series of 18 geodesic regressions help answer this question (Fig. 5b). Each of these regressions fits a geodesic curve to the data as a function of an azimuth (e.g. northing). The maximum R^2 of a geodesic regression is 0.13 (for an azimuth of 30°). A kernel regression, which fits a more complex function to the data, of 30° azimuth has an R^2 value of 0.522.

These low R^2 values suggest that the geographic dependency observed in the equal area and equal volume plots is probably not systematic, and leads to the division of the data into multiple domains (Fig. 5C). The same plotting and regression analysis for each domain suggests that there is no strong geographic dependence.

Within each domain, the data are approximately unimodal and symmetric about that mode, so the mean is an appropriate summary statistic. We use the Fréchet mean, which is the point that minimizes the Fréchet variance (Table 1). The dispersion of the data can be described using the matrix Fisher maximum likelihood estimation. This dispersion measure is not meaningful geologically, but is critical to selecting which inference method is most appropriate (Davis and Titus, 2017). In this case, MCMC simulation is the best behaved method. As a check, bootstrapping has also been done.

MCMC and bootstrapping results for each domain are shown in Fig. 5D and 5E, with 95% credible/confidence ellipsoids. The null hypotheses are that each pair of domains are sampled from populations with the same mean. The credible/confidence regions of domain 2 and 3 overlap appreciably, while the credible/confidence region of domain 1 does not overlap with the other two. The null hypothesis that domains 2 and 3 are sampled from populations with the same mean cannot be rejected. The null hypothesis that domain 1 and domain 3 are sampled from populations with the same mean can be rejected with 95% credibility/confidence. The null hypothesis for domains 1 and 2 can also be rejected with 95% credibility/confidence.

5.2 Summary

There are two first-order conclusions that can be drawn from the statistical analysis of the foliation-lineation pairs near Orofino, ID. First, there are geographic domains; within each domain the data are roughly unimodal and symmetric, and apparent spatial dependencies have consistently low R^2 values. Second, hypothesis testing using bootstrapping and MCMC show that while the difference between orientations of domains 2 and 3 is not statistically significant, domain 1 is significantly different from the other two domains.

These results are consistent with the mapped boundary between the Ahsahka and Woodrat Mountain shear zones as defined by cooling ages. Domains 2 and 3 are along strike of one another, and have been previously interpreted to be part of the Woodrat Mountain shear zone. Domain 1, which is across strike from the other two domains, has been interpreted to be part of the later Ahsahka shear zone. The presence of distinct orientations, confirmed to be statistically significant in this analysis, in rocks with different $^{40}Ar/^{39}Ar$ cooling ages provides further evidence that there were two distinct shear zones, now located adjacent to one another.

6. Discussion

In both the West Mountain and Orofino datasets, a workflow that incorporates statistical analysis leads to interpretations that are tested in an objective way with reported uncertainty—some of which would likely not have been made otherwise. At West Mountain, we show that foliation-lineation pairs are not statistically distinguishable in different geographic areas, and that foliations in the foliation-only dataset are demonstrably different from foliations in the foliation-lineation dataset. In addition, we computed the magnitude of rotation between the northern and southern domains in a way that incorporates the uncertainty about the mean of each domain. In the Orofino area, we show that the difference in orientation on opposite ends of the Dworshak Reservoir could not be accounted for by systematic spatial variations and that the difference in orientation between the southwest shore of the reservoir (domain 1) and the northeast shore (domains 2 and 3) is statistically significant. This division is consistent with other geologic data.

The statistical approach has tangible scientific benefits for structural geology data. Statistical methods help to quantitatively identify spatial tendencies in high-dimensional data. They also can be used to quickly compute basic statistical descriptors of the dataset and uncertainty about the population mean, which helps guide the visualization of data. The uncertainty of the population mean is used to reject (or fail to reject) geologic hypotheses posed as statistical hypotheses. Using this approach to testing hypotheses, geologic interpretations come with quantifiable uncertainties which structural geologists can report in publications. The statistical approach also allows structural geologists to assess the validity of implicit assumptions using the same methods that are used to test geologic hypotheses.

6.1 Identifying spatial tendencies

In our regressions of foliation-lineation data, we treat each data point holistically as a rotation matrix. This approach is an improvement on standard practice in structural geology. The investigation of geographic trends in structural geology data usually involves the decomposition of high-dimensional data such as foliation-lineation pairs into one-dimensional elements. For example, it is common to plot the strike of foliation against the distance from a shear zone, even though each data point is a strike, dip, and rake. These two-dimensional charts have some utility but provide an incomplete view of each data point. For example, the lineation is not independent of the foliation. Best-fit lines and associated R^2 values in these charts are problematic because such regressions should be informed by the other dimensions that comprise each data point. This partial view of the data can lead to false correlations or can fail to reveal correlations entirely. By treating the data holistically during statistical analysis, structural geologists can be more accurate in their identification of spatial dependencies.

Performing regressions is also a critical step when testing that the data are spatially independent. Common techniques for inference about the mean of a population assume this independence. An iterative process of visual inspection (plotting), regression analysis, and division of the data into domains (as performed in the Orofino example) ensures that this assumption is reasonable.

6.2 Uncertainty about the mean

Structural geology datasets are often relatively small and dispersed. This characterization is especially true for field datasets such as the ones described in this paper. Equal-area projection computer programs widely used by structural geologists have built-in measures of mean and dispersion for directional data (e.g. foliations only), but do not yet treat orientation data (e.g. foliation-lineation pairs). For foliation-lineation data, we employ one of two equally valid conceptions of the mean (Davis and Titus, 2017). To quantify how well the mean of the dataset reflects the mean of the population from which the dataset was sampled, bootstrapping and MCMC simulations produce clouds of means from which confidence and credibility regions can be inferred. The confidence/credible region for the mean of a structural geology dataset has two main functions. First, it contextualizes the mean of the dataset—a mean is not particularly useful if the uncertainty about that mean is very large. Second, confidence/credible regions enable comparison with other datasets using hypothesis testing.

6.3 Hypothesis testing

In both examples provided in this paper, an experienced structural geologist would most likely notice differences among some of the domains. Taking a statistical approach, structural geologists can test hypothesized differences in an objective way. In practice, if the 95% confidence/credible regions of two domains do not overlap, the null hypothesis that they are sampled from the same population can be rejected.

Statistical significance is especially important when the difference between two datasets is small or data are dispersed. In the West Mountain example, foliations associated with foliation-only measurements plot in the same general area of the equal area projection as foliations associated with foliation-lineation measurements. While visual inspection may lead a structural geologist to suspect a difference between the two datasets, it is only through statistical hypothesis testing that the geologist can say that this difference is not due to random variation within the same population, at 95% confidence. We are able to rely on this interpretation to ask further questions—such as why foliation-lineation pairs are different from foliation-only data—precisely because we have rejected the null hypothesis that they come from the same population.

6.4 Using hypothesis testing and regressions to assess assumptions

A statistical comparison of the two West Mountain datasets leads to the conclusion that foliations from foliation-only data were likely not sampled from the same population as those from foliation-lineation data. This finding leads us to reject an assumption that Braudy et al. (2017) made that seemed logical. The ability to quickly assess such assumptions is a key advantage of the statistical workflow. In part, this advantage comes from a shift in perspective, because the statistical approach forces the articulation (and thus awareness) of the assumptions we make when analyzing data.

6.5 Better science through statistics

The use of statistics in structural geology may seem onerous, simply another task to complete prior to submitting a manuscript. However, given the examples above we suggest that there are many reasons to adopt this methodology throughout the data analysis and interpretation process. Taken together, the benefits of the statistical approach make it easier to have the scientific integrity Feynman (1974) discussed in his famous essay “Cargo Cult Science”:

> *The first principle is that you must not fool yourself—and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that.*

When conducting fieldwork, many of the hypotheses that we initially formulate are ultimately incorrect. The successful execution of science is the ability to generate and discard hypotheses with relative efficiency. Statistical analysis aids in perhaps the most difficult part of the scientific process: exactly when to discard a hypothesis. Statistical analysis is used by most scientific communities—including field scientists such as ecologists—to facilitate this process. The relatively small size and large dispersion common to structural geology datasets may seem a good excuse not to use statistics. In fact, these characteristics are particularly compelling reasons to incorporate statistical tools into the structural geology workflow. It can be tempting to over-interpret small datasets, and statistics provides a check on what interpretations are permissible given the small sample size.

Finally, the process of science depends on the presentation of data and interpretations to scientific peers. Most data in structural geology papers present data in the form of lower hemisphere projections (e.g. equal area) or other representative documentation, neither of which allows other structural geologists to evaluate or use the dataset effectively. A clear statement of the tested hypotheses and the results would be a useful way to communicate the uncertainty of the data with respect to a specific model. The theory for the types of data that we collect has been addressed by Davis and Titus (2017) and the tools to statistically analyze data are now available. With the addition of specific use cases—as introduced in this contribution—we hope that both the methodology and its utility will be clear and accessible to the structural geology community.

Conclusion

Structural geology, especially structural geology in the field, is a science that benefits from the incorporation of statistical procedures. Field datasets are commonly small, geographically dispersed, and limited to small areas of good outcrop. Further, structural geology data are inherently high-dimensional, meaning that traditional ways of viewing data provide incomplete pictures of the data. The analysis of structural geology data within a statistical framework provides a way for structural geologists to more quantitatively understand and interrogate their data.

In this contribution, two typical structural geology field datasets were analyzed using direction and orientation statistics. In both cases, we employed a workflow in which geologic expertise interacts with statistical protocol to motivate geologically relevant statistical tests (Fig. 1). In this framework, statistics connects the collected dataset to the geologic system through quantitative measures of uncertainty. We find significant utility in adopting such a workflow, particularly for datasets that are small and dispersed. Utilizing a statistical approach allowed us to interpret subtle differences in domains as real through hypothesis testing.

Statistical tools are critical to the future of structural geology. As structural geology datasets become available in open source databases, these statistical tools will be increasingly important. When combining datasets collected by different geologists over the same geographic extent, these tools provide a way to test whether combining datasets is permissible. When examining the same type of geologic feature at thousands of field locations worldwide, these tools provide a way to quantitatively compare geometries.

Acknowledgements

This work was supported by funding from the National Science Foundation EAR 1639748 to B. Tikoff. N. Braudy generously provided the dataset from Braudy et al. (2017). Thank you to V. Chatzaras, N. Garibaldi, R. Williams, A. Jones, C. Bate and the rest of the Structure and Tectonics group at UW-Madison for insightful manuscript comments. We thank reviewers J. Loveless and A. Rotevatn for helpful comments.

References

Allmendinger, R. W. (2017), Stereonet 10.0.0 [Computer software]. Retrieved from http://www.geo.cornell.edu/geology/faculty/RWA/programs/stereonet.html

Allmendinger, R. W., Siron, C. R., & Scott, C. P. (2017). Structural data collection with mobile devices: Accuracy, redundancy, and best practices. *Journal of Structural Geology*, *102*, 98-112. https://doi.org/10.1016/j.jsg.2017.07.011

Armstrong, R. L., W. H. Taubeneck, and P. O. Hales (1977), Rb-Sr and K-Ar geochronometry of Mesozoic granitic rocks and their Sr isotopic composition, Oregon, Washington, and Idaho, *Geological Society of America Bulletin*, *88*(3), 397–411. https://doi.org/10.1130/0016-7606(1977)88<397:RAKGOM>2.0.CO;2

Benford, B., J. Crowley, M. Schmitz, C. J. Northrup, and B. Tikoff (2010), Mesozoic magmatism and deformation in the northern Owyhee Mountains, Idaho: Implications for along-zone variations for the western Idaho shear zone, *Lithosphere*, *2*(2), 93–118. https://doi.org/10.1130/L76.1

Braudy, N., R. M. Gaschnig, D. Wilford, J. D. Vervoort, C. L. Nelson, C. Davidson, M. J. Kahn, and B. Tikoff (2017), Timing and deformation conditions of the western Idaho shear zone, West Mountain, west-central Idaho, *Lithosphere*, *9*(2), 157–183. https://doi.org/10.1130/L519.1

Davidson, G. F. (1990), Cretaceous tectonic history along the Salmon River suture zone near Orofino, Idaho: Metamorphic structural and ^{40}Ar/^{39}Ar thermochronologic constraints (M.Sc. thesis). Location: Oregon State University.

Davis, J. C., and R. J. Sampson (1986), *Statistics and data analysis in geology*, Wiley New York.

Davis, J. R., and S. J. Titus (2017), Modern methods of analysis for three-dimensional orientational data, *Journal of Structural Geology*, *96*, 65–89. https://doi.org/10.1016/j.jsg.2017.01.002

Downs, T. (1972), Orientation statistics. *Biometrika, 59, 665-676.*

Ducharme, G. R. (1985), Bootstrap confidence cones for directional data, *Biometrika*, *72*(3), 637–645. https://doi.org/10.1093/biomet/72.3.637

Efron, B., and R. J. Tibshirani (1994). *An introduction to the bootstrap*. CRC press.

Feynman, R. P. (1974), Cargo cult science, *Engineering and Science*, *37*(7), 10–13.

Fisher, N. I., and P. Hall (1989), Bootstrap confidence regions for directional data, *Journal of the American Statistical Association*, *84*(408), 996–1002. https://doi.org/10.1080/01621459.1989.10478864

Fleck, R. J., and R. E. Criss (1985), Strontium and oxygen isotopic variations in Mesozoic and Tertiary plutons of central Idaho, *Contributions to Mineralogy and Petrology*, *90*(2–3), 291–308. https://doi.org/10.1007/BF00378269

Fleck, R. J., and R. E. Criss (2004), Location, age, and tectonic significance of the Western Idaho Suture Zone (WISZ). *Report: USGS Numbered Series,* 2004-1039, p. 48.

Giorgis, S., W. McClelland, A. Fayon, B. S. Singer, and B. Tikoff (2008), Timing of deformation and exhumation in the western Idaho shear zone, McCall, Idaho, *Geological Society of America Bulletin*, *120*(9–10), 1119–1133. https://doi.org/10.1130/B26291.1

Giorgis, S., Z. D. Michels, L. Dair, N. Braudy, and B. Tikoff (2017), Kinematic and vorticity analyses of the western Idaho shear zone, USA, *Lithosphere*, *9*(2), *223–234*. https://doi.org/10.1130/L518.1

Jupp, P. E., and K. V Mardia (1989), A unified view of the theory of directional statistics, 1975-1988, *International Statistical Review/Revue Internationale de Statistique*, *57*(3), 261–294. https://doi.org/10.2307/1403799

Lewis, R. S., J. H. Bush, R. F. Burmester, J. D. Kauffman, D. L. Garwood, P. E. Myers, and K. L. Othberg (2005). Geologic map of the Potlatch 30× 60 minute quadrangle, Idaho. *Idaho Geological Survey Geological Map*, *41*.

Lewis, R. S., P. K. Link, L. R. Stanford, and S. P. Long (2012). Geologic map of Idaho. *Idaho Geological Survey Map*, *9*.

Lewis, R. S., K. L. Schmidt, R. M. Gaschnig, T. A. LaMaskin, K. Lund, K. D. Gray, B. Tikoff, T. Stetson-Lee, and N. Moore (2014), Hells Canyon to the Bitterroot front: A transect from the accretionary margin eastward across the Idaho batholith, *Geological Society of America Field Guides*, *37*, 1–50.

Mahalanobis, P. C. (1936), On the generalized distance in statistics, *Proceedings of the National Institute of Sciences of India, 1936*, 49–55.

Manduca, C. A., L. T. Silver, and H. P. Taylor (1992), $^{87}Sr/^{86}Sr$ and $^{18}O/^{16}O$ isotopic systematics and geochemistry of granitoid plutons across a steeply-dipping boundary between contrasting lithospheric blocks in western Idaho, *Contributions to Mineralogy and Petrology*, *109*(3), 355–372. https://doi.org/10.1007/BF00283324

Manduca, C. A., M. A. Kuntz, and L. T. Silver, L. T. (1993). Emplacement and deformation history of the western margin of the Idaho batholith near McCall, Idaho: Influence of a major terrane boundary. *Geological Society of America Bulletin, 105*(6), 749-765. https://doi.org/10.1130/0016-7606(1993)105%3C0749:EADHOT%3E2.3.CO;2

Mardia, K. V, and P. E. Jupp (2000), *Directional statistics*, John Wiley & Sons.

Marrett, R., and R. W. Allmendinger (1990), Kinematic analysis of fault-slip data, *Journal of structural geology*, *12*(8), 973–986. https://doi.org/10.1016/0191-8141(90)90093-E

McClelland, W. C., and J. S. Oldow (2004), Displacement transfer between thick-and thin-skinned décollement systems in the central North American Cordillera, *Geological Society, London, Special Publications*, *227*(1), 177–195. https://doi.org/10.1144/GSL.SP.2004.227.01.10

McClelland, W. C., and J. S. Oldow (2007). Late Cretaceous truncation of the western Idaho shear zone in the central North American Cordillera. *Geology, 35*(8), 723-726. https://doi.org/10.1130/G23623A.1

Michels, Z. D., S. C. Kruckenberg, J. R. Davis, and B. Tikoff (2015). Determining vorticity axes from grain-scale dispersion of crystallographic orientations. *Geology, 43*(9), 803-806. https://doi.org/10.1130/G36868.1

Novakova, L., and T. L. Pavlis (2017). Assessment of the precision of smart phones and tablets for measurement of planar orientations: A case study. *Journal of Structural Geology*, *97*, 93-103. https://doi.org/10.1016/j.jsg.2017.02.015

Schmidt, K. L., R. S. Lewis, J. D. Vervoort, T. A. Stetson-Lee, Z. D. Michels, and B. Tikoff (2017), Tectonic evolution of the Syringa embayment in the central North American Cordilleran accretionary boundary, *Lithosphere*, *9*(2), 184–204. https://doi.org/10.1130/L545.1

Stetson-Lee, T. A. (2015), Using Kinematics and Orientational Statistics to Interpret Deformational Events: Separating the Ahsahka and Dent Shear Zones Near Orofino, ID (M.Sc. thesis). Location: University of Wisconsin--Madison.

Strayer IV, L. M., D. W. Hyndman, J. W. Sears, and P. E. Myers (1989). Direction and shear sense during suturing of the Seven Devils-Wallowa terrane against North America in western Idaho. *Geology, 17*(11), 1025-1028. https://doi.org/10.1130/0091-7613(1989)017%3C1025:DASSDS%3E2.3.CO;2

Tikoff, B., P. Kelso, C. Manduca, M. J. Markley, and J. Gillaspy (2001). Lithospheric and crustal reactivation of an ancient plate boundary: The assembly and disassembly of the Salmon

River suture zone, Idaho, USA. *Geological Society, London, Special Publications*, *186*(1), 213-231. https://doi.org/10.1144/GSL.SP.2001.186.01.13

Tsagris, M., and G. Athineou (2016). Directional: Directional Statistics. *R package version 1.8.* http://CRAN.R-project.org/package=Directional

Vollmer, F. W. (2017), Software for the quantification, error analysis, and visualization of strain and fold geometry in undergraduate field and structural geology laboratory experiences, in *Geological Society of America Abstracts with Programs*, *49*(2). https://doi.org/10.1130/abs/2017NE-291027

Wellner, J. A. (1979), Permutation tests for directional data, *The Annals of Statistics*, *7*(5), 929–943. http://www.jstor.org/stable/2958664

Yonkee, W. A., and A. B. Weil (2015), Tectonic evolution of the Sevier and Laramide belts within the North American Cordillera orogenic system, *Earth-Science Reviews*, *150*, 531–593. https://doi.org/10.1016/j.earscirev.2015.08.001

Figure Captions

Figure 1. Schematic diagram of a structural geology workflow that takes advantage of statistical tools to aid interpretations of the geologic system. The grey box surrounds the statistical component of the workflow and is a simplification of the statistical flowchart from Davis and Titus (2017). Grey arrows denote steps that involve regressions. Thicker arrows represent paths that are taken in the statistical analysis of the Orofino dataset in this paper. The structural geologist begins with an incomplete representation of the geologic system (the dataset). After visualizing the data, two simultaneous processes begin—the generation of geologic hypotheses

and predictive models as well as a statistical protocol that should be done on any dataset. Importantly, all interpretations of the geologic system run through the grey statistical box.

Figure 2. Simplified geologic map and overview of data from the West Mountain, ID area of the Late Cretaceous western Idaho shear zone published in Braudy et al. (2017). **A)** Geologic units of the western Idaho shear zone (Red—Muir Creek orthogneiss, Purple—Sage Hen orthogneiss, Magenta—Payette River Tonalite) superimposed on a hill shade model of topography. The Muir Creek orthogneiss was the focus of the structural study in Braudy et al. (2017). Inset map shows the location of the field area on the Western Idaho Shear Zone (WISZ), shown by the red line. **B)** A cutout of the Muir Creek orthogneiss with hill shade model of topography, showing the geographic locations and symbols of foliation-lineation data (left) and foliation-only data (right). There are 148 foliation-only measurements and 129 foliation-lineation pairs. **C)** Equal area nets with data for foliation-only (left) and foliation-lineation datasets (middle). Also shown is an equal volume plot (right) (Davis and Titus, 2017), in which each line-plane pair is represented by a single point, and which shows four symmetric copies of the data. All plots are color-coded by the geographic domains used by Braudy et al. (2017) (Red—northern, Green—central, Blue—southern). Map modified from Braudy et al. (2017).

Figure 3. Summary of the statistical analyses for the West Mountain field fabrics dataset, the location for which is shown in Figure 2. **A)** An analysis of the claim from Braudy et al. (2017) that there is a 20° rotation between the northern and southern domains: Top, a lower hemisphere equal area projection (with zoomed-in cutout) with the 95% confidence regions for the mean of foliation-only data in each of the three domains (Red—northern, Green—central, Blue—

southern) as determined from bootstrapping; Middle, a histogram of angular distances between bootstrap iterations of the northern and southern domains; Bottom, a lower hemisphere equal area net projection visualization of the rotation computed from the bootstrapped angular distance and corresponding rotation axes. **B)** A series of two-sample hypothesis tests plotted on equal volume plots (with zoomed-in cutouts). Both bootstrapping and 95% confidence ellipsoids as well as Markov chain Monte Carlo (MCMC) mean probability clouds and their 95% credible ellipsoids are used to compare each pair of domains (Black—northern, Orange—central, Blue—southern). **C)** A comparison of 95% confidence ellipses from bootstrapping foliations. Foliations from foliation-lineation data are compared with those from foliation-only data within each domain: Colors are the same as in (A).

Figure 4. Simplified geologic map of the Orofino area, with the foliation-lineation dataset superimposed. A cutout map of Idaho shows the location of the Orofino field area. The red line shows the location of that Ahsahka shear zone, interpreted as part of the western Idaho shear zone. Exposure of sheared Late Cretaceous basement below the Miocene Columbia River basalts is limited to the shoreline of Dworshak reservoir, where all foliation-lineation pairs were measured. An interpretation of the boundary between the Woodrat Mountain and Ahsahka shear zones is shown. Modified from Lewis et al. (2005) and Lewis et al. (2012).

Figure 5. Summary of statistical analysis for the Orofino, ID area foliation-lineation dataset. **A)** Two different plots of the foliation-lineation data colored by kilometers north in UTM: Left, an equal-area plot with lineations (squares) and foliation poles (circles), each with 2σ, 6σ, 10σ, 14σ, and 18σ Kamb contours; Right, an equal volume plot after Davis and Titus (2017) with

translucent 2σ Kamb contours. Each point in the equal volume plot is a foliation-lineation pair represented as a rotation from a reference plane-line pair. Note that there are four copies of the dataset due to four-fold symmetry of such data (See Davis and Titus (2017) for more information). **B)** A series of 18 geodesic regressions testing geographic variation along specific azimuths. Each solid dot is a regression with a corresponding *p*-value based on 100 permutations (open circle). **C)** The geologic map from Figure 4 superimposed with the domains used in this statistical analysis. **D)** A series of two-sample hypothesis tests plotted on equal volume plots (with zoomed-in cutouts). MCMC mean probability clouds and their 95% credible regions as well as bootstrapped mean clouds and their 95% confidence region are used to compare each pair of domains (Black—domain 1, Orange—domain 2, Blue—domain 3). **E)** A lower-hemisphere, equal-area projection showing the results of the MCMC analysis. It can be seen that both the foliation and lineation are different for Domain 1.

Table 1. Two conceptions of the mean strike, dip, and rake for the three domains in the Ahsahka segment of the western Idaho shear zone. Strike/dip is in right hand rule.