Reinforcing Quantitative Skills with Applied Research on Tombstone-Weathering Rates

Sheila Marlene Roberts
Environmental Sciences Department
Western Montana College of the University of Montana
Dillon, Montana 59725
s_roberts@wmc.edu

ABSTRACT

Students in a surficial-processes class helped design and completed a research project to measure the weathering rates of marble tombstones at two locations in southwestern Montana. Measuring the rates of rapid geomorphic processes emphasizes the quantitative aspects of research project design, data collection, and analysis. The project reinforced students' recently acquired knowledge of statistics (from a required math class) as they used it, at their own initiative, to solve a geologic problem. When the field data were more complex than they expected, students demonstrated for themselves that mathematical analysis could give meaning to their data.

Keywords: Education – geoscience; education – undergraduate; miscellaneous and mathematical geology; surficial geology – geomorphology.

INTRODUCTION

Quantitative skills are reinforced by research experiences in which students are involved in identifying a problem that can be addressed quantitatively, designing the research approach, gathering data, and interpreting the results by statistical analysis. Geoscience undergraduates at Western Montana College are required to take a statistics class as part of their Environmental Sciences degree. Integrating statistics into geology class research projects helps students develop an understanding of when and how to use statistics to solve scientific problems.

I teach our surficial-processes course with an emphasis on the rates of geomorphic change and expect students to complete an applied-research project in that area. The project is about one-third of their grade for the class. It is a team effort, and students determine for themselves how the work is to be divided up and accomplished. Most students take the required statistics course in their second year and surficial processes in their third or fourth year, so I can build teams in which at least some members have statistical skills. I offer myself as part of the data-gathering team and act as a project advisor and discussion leader when they request my assistance. We are fortunate to have a backyard full of potential projects in geomorphic process rates, but the research project described here could be duplicated almost anywhere.

The type of project described here may be limited by class size. Our degree program is new and upper-division geology classes are still small (5 to 15 students). However, there is a strong applied field research emphasis in all our environmental sciences classes, and we have found that we can do open-ended field projects with groups of up to about 25 students. It may also require somewhat nontraditional scheduling. Surficial processes meets for one hour on one day a week (lectures and exams) and three hours on another day (some lectures, discussions, labs, field work). With advance notice, students will agree to extend some field days to four or five hours.

This paper describes a geology class research project where students developed a deeper conceptual understanding of, and a greater ability to apply, statistical skills they had learned in a required mathematics class. The narrative is largely chronological because describing how the project developed is key to explaining the process by which students integrate their research and quantitative skills.

PROJECT DESIGN

Designing a project that will produce interpretable quantitative data is the first aspect of developing quantitative skills through applied research. I set the basic constraints on the design of the research project. To help overcome students' inexperience, I provide a list of possible projects but also give them the opportunity to pick another topic. In 1999, I offered the idea of measuring the weathering rates of tombstones, which is at least as old as the first published scientific paper about it in English (Geikie, 1880) and has these advantages related to my constraints:

1) Accessibility – cemeteries are accessible in all seasons;
2) Datability – tombstone dates provide a known time frame;
3) Measurable change – previous work suggests that measurable change occurs in most places within 100 years;
4) Possibility of creating a large data set in a short time – hundreds of measurements can be quickly acquired;
5) Available research methods – previous workers have published several methods for measuring the weathering rates of tombstones;
6) Scientific value – there is a need for more quantitative information on natural rock-weathering rates, for example, to predict the life of and protect public buildings and monuments and to understand global cycles, especially carbon in this case (tombstone-weathering rates for semiarid and arid climates are not readily available); and

7) Low cost – costs include local transportation and calipers (which we borrowed).

After being presented with their options, students chose the tombstone work. Following a directed in-class discussion, they reviewed the literature based on questions of their own construction: What is the extent of previous work on this subject? What basic assumptions were made? What methods have been used? What weathering rates have been determined elsewhere? What kinds of problems have previous studies encountered? In the second week, students devised a preliminary research plan based on their reading.

Next, we visited the local (Dillon, Montana) cemetery, which contains tombstones dating back to the 1870s. Students were immediately confronted with unanticipated variables. Of the two major stone types (marble and granite), only the marble ones had weathered noticeably, even after 100 years. There were many different tombstone shapes, and some stones had chipped or spalled. The sprinkler system and protecting vegetation compromised the climate description to some extent.

After considering these and other complexities, students discussed what was required for a quantitative study and narrowed their project based on the potential for gathering the largest number of comparable measurements with the fewest uncontrolled variables. They determined to measure all upright, relatively pure, marble (calcite or dolomite) slab tombstones that lacked significant spalling or breakage. The decision to measure all tombstones that fit these criteria committed the students to more work but avoided establishing a statistically random sample and determining how many measurements were enough. Two of the stronger students brought up the question of "the sample" and led the class discussion to that conclusion.

Dillon is a small rural community in a fairly pristine semiarid setting. We also visited three cemeteries in Butte, an old mining and smelting city of similar climate. A quick visual reconnaissance demonstrated that there had been more extreme weathering of marble tombstones in Butte, which also dated back to the 1870s. In Butte, some tombstones were weathered beyond legibility and some did not have dates, but the cemetery caretaker provided old plot records with that information.

Finally, students determined that the (quantifiable) testable hypotheses of their research would be: (1) thickness loss is a direct and relatively consistent function of exposure time and (2) stones in Dillon have weathered at a measurably slower rate than stones in Butte.

METHODS AND MEASUREMENTS

Determining exactly what to measure and how to measure it is the second aspect of quantitative skill development through applied research. The literature review revealed several possible types of measurements, all based on loss of surface material due to weathering. The students rejected degradation of

Figure 1. Students at work measuring marble tombstones in the Dillon, MT cemetery.

inscription legibility, as described by Rahn (1971) and Meierding (1993), because of the level of subjectivity involved in taking the "measurements." Measurement of inscription-depth reduction (Matthias, 1967) was rejected because the original depth of inscriptions was unknown. After assessing the available materials and research tools, the students chose to compare the top thickness to the base thickness of tombstones to determine surface recession, a measure of local weathering rate described in Feddema and Meierding (1986). This method assumes that the top of the tombstone is exposed to far more fresh acidic precipitation than the base, which receives mostly neutralized surface runoff from the upper part of the stone. If acid precipitation is the primary weathering agent, the top should exhibit maximum loss of thickness and the base should be virtually unweathered, so that a comparison of the two would give a theoretical initial and final state. In fact, the lower few centimeters of the most visibly weathered older tombstones still retained polished surfaces and angular corners, while the tops were rough and rounded.

Some other critical assumptions in this method that could be tested in this study included:

- Initial thicknesses of top and bottom were identical (to help verify this, one student volunteered to try to find thickness-control information from tombstone suppliers);
- Weathering is relatively even across the top; and
- There has been measurably significant thinning during the time span available.

Assumptions that students determined to be important but outside the scope of this study included:

- The primary measurable effect of weathering is thinning of the stone;
- Thickness loss is mostly by in-place dissolution of crystals; and
- The tombstones were placed in the cemetery on or very near the inscribed date of death.

In the field, students worked in teams of two or three (Figure 1). They measured tombstone thicknesses with dial calipers accurate to 0.001 inch (we did not have metric calipers). The five thickness measurements of each tombstone included one on either side of the base and three across the top (at the right and left edges and in the middle). This allowed them to compare the average top measurement with the average base measurement.

In a few visits to the cemeteries, the class measured virtually all available tombstones and gathered a total of more than 400 measurements. The most problematical part of the fieldwork was deciding when the stone had too many mineral impurities (mostly amphiboles, pyroxenes, and micas) to be included. The students did not establish a satisfactory quantifiable cutoff, but they avoided taking measurements directly on visible mineral impurities and later decided to eliminate some measurements where their notes indicated "high" amounts of impurities.

I could have introduced the question of precision early in the study but chose to wait until it occurred to the students. Eventually, the loss of a page of data required them to remeasure some tombstones. One student suggested that the original team should do the remeasurements. Another wondered if that would make a difference. Immediately, they realized the importance of determining the precision of their measurements, a concept previously covered in chemistry and statistics classes, and now strongly reinforced. They then remeasured tombstones at three different cemeteries to check precision, collecting data on the same and different tombstones measured by the same student and by different students.

Students created their own data-collection sheets and entered the data into a computer spreadsheet. Most of the students were unfamiliar with the program (*Excel*), but they learned the basics quickly in class and worked on it together outside class time. A warning: students working on the same project, using different computer labs and home computers, can have trouble maintaining records. At some point, the collected data with all its calculations and graphs justifiably acquired the folder name "Reese's mess" (for the student who was keeping track of it all for the class). I periodically downloaded everything onto my computer as a safeguard against loss.

ANALYSIS

Data analysis was the final aspect of this project, and the most central one for reinforcing statistical skills. Somewhat surprisingly, the students collected most of their data before they turned to analysis. The lag between data collection and analysis may have been partly a scheduling problem because for about a month we spent the long day on project field trips and the short day on other subjects. It may also have been partly their unfamiliarity with the options for computerized analyses. Finally, I had not pressed them for results. With future classes, I may ask more questions if analysis seems to be lagging, but I also think allowing students a little stumbling space is useful. Their experience in this class should help them appreciate the value of ongoing assessment of results during data collection.

The first and simplest proposed analysis was to extract and compare two numbers that would presumably be the mean surface recession rate for each site. They subtracted the top measurement average from the base measurement average for each tombstone at the two sites and averaged those differences, expecting positive values (bases thicker than tops) that were satisfyingly higher at Butte than at Dillon. They did this calculation on the spreadsheet, writing the formula with some help from me, and found that both numbers were zero, to three decimal points. Reviewing the data, they recognized three problems:

1) Some tops were thicker than bases – the exact opposite of the expected results.
2) The first analytical approach was biased by the large number of young stones with negligible weathering. They discussed and discarded a suggestion to use only stones older than some cutoff date and then decided to take a look at the problem by reordering the data and graphing thickness difference against the age of the tombstone.
3) The point graph comparing the age of tombstones with differences between average top and base measurements (Figure 2) produced a clear visual image of a surprisingly wide range of thickness differences for individual tombstones instead of the anticipated neat linear array.

Students spent a week contemplating the unexpected aspects of their results (described above), during which they all displayed considerable frustration. However, they remained serious about the results and showed great integrity in dealing with their data and analyses. Two of them temporarily gave up, but they were drawn back into the discussion by the others. One of the first considerations was to analyze the precision data, which had been collected at the end of a long field day and then neglected. The measurement error was almost an order of magnitude less than the average difference between top and bottom measurements. Next, they considered the possibility that some of the data were flawed. They rechecked all the typed numbers against original field-data sheets and found several errors. The field notes also uncovered some tombstone data that should have been eliminated because of mineralogical variations. But these corrections did not substantially alter the results.

The most problematical aspect of the data was a large set of tombstones in both cemeteries in which the tops were thicker than the bottoms (#1 above). The precision analysis suggested this was not measurement error. The students observed that the number of tombstones with negative differences (base to top) was two or three times less than tombstones with positive differences, which supported the conclusion that tops weather faster than bases (Figure 2). They discussed and rejected the idea of simply eliminating all the negative values. Students critiqued their assumptions and decided either (a) the assumption that weathering should be greater at the tops than at the bottoms was wrong or (b) the assumption of original identical thickness was wrong. One student had unsuccessfully

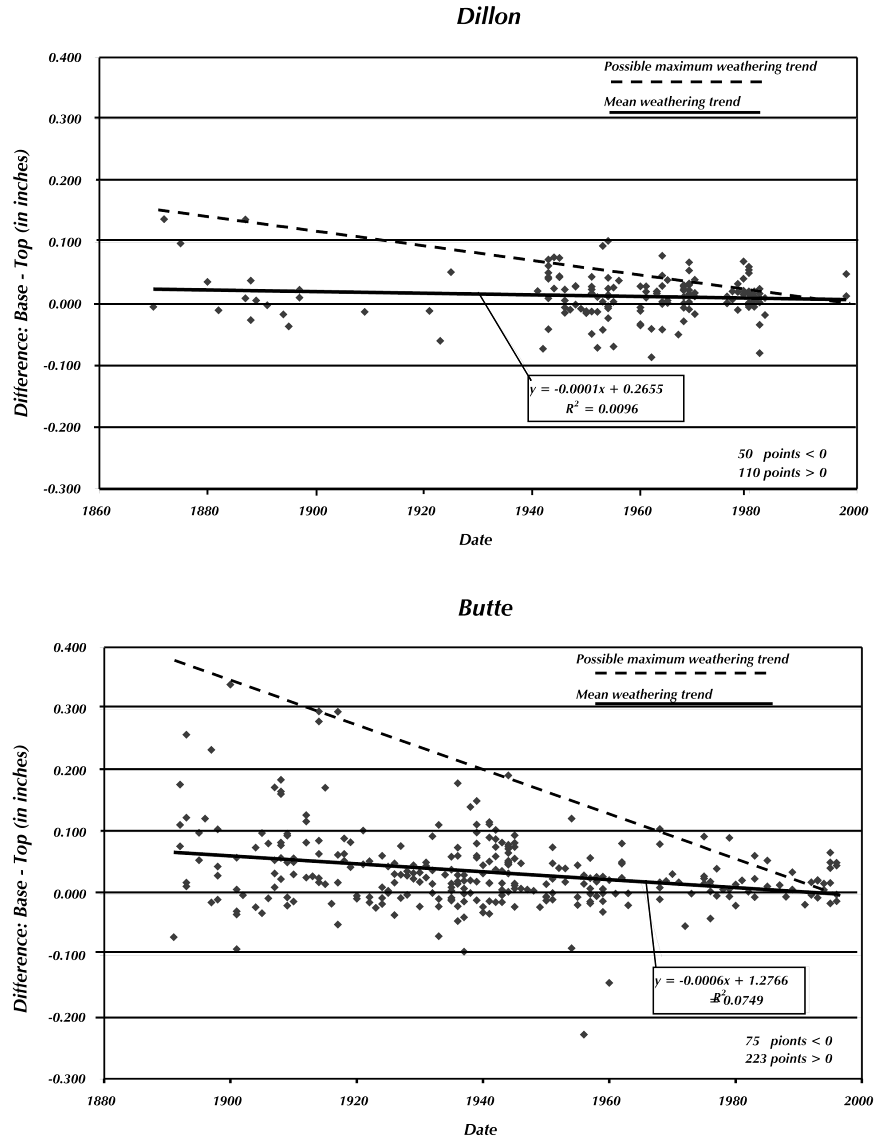

Figure 2. Point graphs of the difference between base and top thicknesses of marble tombstones as a function of the date on the tombstone for the Dillon (top) and Butte (bottom) cemeteries. Graphs also show the linear regression (mean weathering trend) and r-squared values for the data. The number of tombstones that showed the expected trend of base thicker than top (difference >0) and the unexpected number of tombstones with top thicker than base (difference <0) appears at the lower right. These two graphs were produced by students and demonstrate some of the analyses they performed.

attempted to contact suppliers of the tombstones. He did find that contractors for veterans' tombstones are requested to maintain a tolerance of 1/8 inch (0.375 mm) – well within the range of our measured differences! In response to this information, the students decided that it was reasonable to assume the placement of thick end up or down would be random. If so, there would be an equal exaggeration of negative and positive differences, which would approximately cancel each other in the regression analysis.

Students were unprepared for the wide range of thickness differences that they found for individual tombstones because the existing published literature did not report the raw data or its variability. Students observed that the range was exaggerated by original tombstone thickness differences and concluded that they did not know how to eliminate that effect. They required my encouragement to consider that a natural system could produce a wide range about the mean and that their data might contain some interesting new information.

Students reexamined the corrected data to see how it could be applied to test their hypotheses.

1. Some level of covariance of the amount of surface weathering with length of exposure was obvious from the point graph. However, it was also clear that variance from the regression line was very high. Students calculated the range and standard deviation for the graph of each site to quantify variability. Calculated r-squared values for the samples (Figure 2) quantified the low value of the regression line as a predictor of particular tombstone thickness differences as a function of age. In an attempt to find better correlation, they averaged thickness differences in Butte by decade and fit linear and exponential regression lines (not shown in figures). Both of these produced r-squared values near 0.8, providing stronger evidence of covariance at the level of decadal averages. This was one instance where students could see that their statistical analysis actually added meaning to the data.

2. The hypothesis that stones in Dillon have weathered at a measurably slower rate than stones in Butte was confirmed. The linear regression line provided a visual representation of central tendency, and the resulting slope was steeper for Butte than for Dillon, as expected. It also provided an estimate of surface recession rate. Calculated mean surface recession in Dillon was 0.010 inches/100 years (0.25 mm), while for Butte the data suggested an average weathering rate about five times higher, at 0.050 inches/100 years (1.27 mm). At my suggestion, students estimated a maximum weathering trend by simply drawing a line on the upper points of each graph (Figure 2). They recognized that no additional mathematical analysis was justified because the line is skewed by an unknown number of tombstones with originally thinner tops.

DISCUSSION

The value of research as an educational experience is not something any scientist questions. These aspects of measuring tombstone weathering rates fostered the students' conceptual development and are relevant to the theme of reinforcing quantitative skills:

1) Students confronted some central problems of designing and carrying out a field study to gather meaningful quantitative data, for example:

   How do natural processes work in ways that are measurable and how do we measure them?

   How do we address systematic error in measurements?

   How do we identify and respond to uncontrollable (random) error in measurements? (for example, the problem of original tombstone thickness differences)

2) The analysis progressed in a logical way from simple to more sophisticated statistical approaches, and students understood why each step was necessary to find meaning in the data they collected. They discovered that quantitative analysis is part of the creative process of doing research.
3) Students also confronted some of the temptations to manipulate data inappropriately. In the process, they protected the integrity of their data and analysis, at least partly because they felt ownership in it and the process.
4) The research resulted in a poster session at the 1999 Geological Society of America meeting in Denver (Roberts and others, 1999) and a scientific paper (in preparation). Two of the students attended the poster session, and I was impressed by the clarity and confidence with which they explained their work.
5) The quantitative approach also helped identify problems that remain unanswered by this research but might be productive areas for further investigation.

ACKNOWLEDGMENTS

Kyle Flynn, Reese Gannon, Bryce Nelson, Rowdy Brooks, and Ryan Anderson were the students in the surficial-processes class; they had enough enthusiasm for this project to make it work, which is always the key factor. Suggestions from editors Heather Macdonald and LeeAnn Srogi, and several unnamed reviewers, improved the manuscript.

REFERENCES CITED

Feddema, J., and Meierding, T., 1986, Marble weathering and air pollution in Philadelphia: Atmospheric Environment, v. 21, p. 143-157.

Geikie, F.R.S., 1880, Rock-weathering, as illustrated in Edinburgh churchyards: Proceedings of the Royal Society of Edinburgh, Session 1879-80, p. 518-533.

Matthias, George, 1967, Weathering rates of Portland Arkose tombstones: Journal of Geological Education, v. 15, p. 140-144.

Meierding, T.C., 1993, Inscription legibility method for estimating rock weathering rates: Geomorphology, v. 6, p. 273-286.

Rahn, Perry, 1971, The weathering of tombstones and its relationship to the topography of New England: Journal of Geological Education, v. 19, p.112-118.

Roberts, Sheila, Flynn, Kyle, Gannon, Reese, Nelson, Bryce, Brooks, Roderick, and Anderson, Ryan, 1999, Measuring weathering rates of marble tombstones in Southwest Montana: Geological Society of America 1999 Annual Meeting Abstracts with Programs, v. 31, no. 7, p. A-348.

Food for Thought

In hunter-gatherer, pre-agricultural times, the human life expectancy was about 20 to 30 years. That's also what it was in Western Europe in Late Roman and in Medieval times. It didn't rise to 40 years until around the year 1870. It reached 50 in 1915, 60 in 1930, 70 in 1955, and is today approaching 80 (a little more for women, a little less for men). The rest of the world is retracing the European increment in longevity. What is the cause of this stunning, unprecedented, humanitarian transition? The germ theory of disease, public health measures, medicines and medical technology. Longevity is perhaps the best single measure of the physical quality of life. (If you're dead, there's little you can do to be happy.) This is a precious offering from science to humanity – nothing less than the gift of life.

Carl Sagan, 1995, *The demon-haunted world – science as a candle in the dark*: New York, Random House, 460 p. (from p. 10).