Minimum Acceptance Criteria for a Thesis Related to Geostatistics

Clayton V. Deutsch

This short note addresses the presentation of research results in a written thesis document. This is not a typical technical paper of CCG, but may prove useful to students and industry sponsors. Although some of these criteria are universal, my audience is limited to graduate students under my supervision. The details of research and the suitability of the research results for a thesis are not addressed here. My concern is the actual document being presented in partial fulfillment for a graduate degree. A thesis is the personal statement of the student; however, there are general requirements and accepted standards. Every thesis will be judged by a group of experienced professors. There are many resources related to writing and presenting research results. The modest goal of this short note is to present my view of minimum requirements. If these requirements are not met, then I have an excuse to stop reading.

A thesis is owned by the student. The thesis is also a reflection on the advisor. The advisor must review the thesis and accept the results. According to University guidelines, the supervisory committee must all agree that the thesis is ready to go forward to a defense. In practice, most committee members defer to the advisor, but everyone signing off on the thesis must be satisfied that the thesis adheres to a high standard of scholarship.

Students funded by the Centre for Computational Geostatistics (CCG) should read this document and treat it like a checklist of the minimum criteria for a thesis. Although students are the author of their own theses, they normally want the approval and signatures of three or more other professors. They will get those signatures more readily when their thesis meets these criteria.

All students have to receive some training in ethics and intellectual integrity. The focus of this note is presentation, but the importance of academic honesty and ethical representation of research cannot be underestimated. An entire thesis can be tainted by a single violation.

- □ My thesis adheres to the highest standards of intellectual integrity and honesty.

Content

The slickest presentation, most beautiful format and elegant style cannot overcome a lack of content; crappy research cannot make a good thesis. The content must be compelling and demonstrate independent scholarly research. The student is expected to have undertaken reputable research and have some original results to present. Students in geostatistics do a lot of good research; however, many students do not know when to shift their focus from experimental research to writing their thesis. Students must trust their advisor. If an experienced advisor suggests that there is enough content to start writing – then start writing. Additional work will be required as the thesis takes shape, but that is rarely a major problem. Some additional research and support will be a welcome break from writing.

- □ My advisor told me to start writing.

All students were taught about the scientific method. You have likely forgotten about it. Scientists and the philosophers of science agree that there is a systematic method to investigate phenomena, acquire new knowledge, or investigate established principles. You should be familiar with the principles of the scientific method. In particular, the principles of reasoning, collecting data, formulating and testing hypotheses. There is an expectation that your process of conducting research must be objective. Your data must be shared. Your methodology must be transparent and open to the scrutiny of other researchers.

- □ My thesis is based on the principles of the Scientific Method

Few research subjects are completely new. Most theses present a new solution to a problem that has already been considered by many previous researchers. The new solution is based on a different approach, improved computing technology, algorithms from other disciplines and so on. A thesis must explain the background to the problem being addressed and fairly appraise competing techniques and solutions that have been proposed by others. This is required to adequately defend your work.

□ All relevant previous work and competing techniques have been reviewed.

Research in numerical modeling requires many choices, assumptions, and parameter selections. These must be explained and justified relative to all reasonable alternatives. You have the burden of proof. Some judgment is required. You cannot explain everything, but neither can you be entirely arbitrary and expect the reader to accept what you say.

□ Every consequential choice made in my thesis is defendable and explained.

Experts in numerical modeling will expect all results to be checked to the maximum extent possible. It is unlikely that any expert will accept that your results are better because they ***should be.*** A numerical model that supposedly accounts for more data or that is supposed more sophisticated cannot stand on a simple claim of superiority. It is necessary to compare the results to competing techniques, show improved results with cross validation or the jackknife, and show how your results translate to improved decisions. An improved numerical model may or may not lead to a better decision in the end. To the maximum extent possible, you must validate, check, and accumulate evidence that your methodology works.

□ All of my developments have been compared to competing techniques.

Results must be explained and analyzed. Too many students show results and incompletely describe why the results are important. Perhaps the results are understood by the student and they simply do not explain them. Perhaps the results are not completely understood. Any results that are worthy of appearing in your thesis should be understood and explained.

□ I explain the theoretical and practical significance of all of the results that I show.

There is an expectation that scientific results are reproducible. Some professors would insist that all theses be reproducible. This is tricky given the complexity of modern software, large datasets and complex processing with many parameters. Nevertheless, we should not give up on the goal of reproducibility. Even if people could not reproduce every numerical result, the conclusions should be reproducible with similar numerical analysis.

□ To the largest extent possible, all results can be reproduced by someone else.

No methodology is perfect. An important aspect of scholarly work is an honest assessment of the limitations and range of applicability of the proposed methodology. This assessment should come out of the points mentioned above, but also by conscious choice.

□ The limitations and range of applicability of my research are clearly stated.

Abstract and Introduction

The abstract and introduction should be understandable by a wide audience. Avoid jargon at the beginning and briefly explain jargon as the thesis builds. The final defense will take place with a group of professors in the general subject area, but not all experts in geostatistics. The untainted view of an external expert is refreshing and important. The student is not writing to their advisor. They are writing to a larger scientific audience to document their independent scholarly research. The context of the research within resources engineering must be explained. The document should be understandable to the largest possible audience.

The abstract is particularly important. Each student should read *Scrutiny of the Abstract* by Landes and apply the important points from that document. In addition to the suggestions in that document, the abstract should be understandable to anybody with an undergraduate degree in engineering or related discipline. It is wrong to try and explain all of the details too soon and risk readers missing the entire purpose of the thesis.

□ I have read and tried to apply the principles in *Scrutiny of the Abstract*.

□ My abstract is understandable to a wide audience.

Readers form an opinion of the thesis based on their first impression that is largely due to the introductory chapter. The introduction is the ultimate research proposal, but written after the research is done. It answers three important questions. What problem am I considering? What have others done on this problem? What am I going to do on this problem? The beginning of the introduction should be

understandable to a wide audience. The use of jargon and specialized vocabulary builds slowly and only as necessary. Acronyms are defined upon first usage and used sparingly. If an acronym is used rarely in the thesis, then it should be spelled out each time. The Introduction should leave no doubt in the reader's mind as to why the research is important and what research is being attempted.

□ My Introduction tells people exactly what I am doing and why it is important.

A clear thesis statement must be made clear in the Introduction. The thesis statement should be a concise *scholarly claim*. The scholarly claim relates to the purpose of the entire document. Students should understand exactly what constitutes a *thesis*. A thesis is not a diary or journal of research activities. It is the presentation of a scholarly claim and the presentation of research to support or defend the scholarly claim.

□ I have a clear thesis statement and it is presented in my Introduction.

Writing Style

Style is personal; however, there are accepted standards. These standards have evolved over years and hundreds of theses. It is unlikely that you know better than your advisor or the hundreds that have come before you. Even if you do, you must pick your battles. There is much room to express your creativity and individuality within the framework of accepted standards. The little book *Elements of Style* by Strunk and White is the single most important English reference book. It is short and to the point. The latter half of this book relates to style. Read it.

□ I've read *Elements of Style* and am trying to apply the principles.

A thesis document explains many complex ideas and attempts to tie them together within the framework of a scholarly claim. Clear logical transitions are essential. It is very annoying for a reader to encounter a random jump or disconnected step in the logical flow. All of the text must build consistently toward the logical scientific support of your thesis. Consider knowledge as elevation. Your challenge is to guide readers to a higher elevation. A good guide does not lead clients up cliffs and down canyons and along tortuous paths. Readers appreciate your efforts to guide them along a route that is not too steep, long or tortuous.

□ The sentences and paragraphs in my thesis logically follow one another.

A thesis is not a mystery novel. It is wrong to hold back information or try to surprise the reader. The reader should know the end of the story from the beginning. The thesis document is required to support your claim, but it should lead to no surprises that are not stated at the beginning. Ideally, readers will accept your claim, but they will be interested in seeing how you have reached your conclusions.

□ I've clearly explained in simple terms the point of my thesis at the beginning.

Short simple sentences are the most effective. Long convoluted sentences cause the reader to stop reading. Long tortuous text is the result of unclear thinking. Mark Twain wrote a long letter to a friend and apologized saying that he did not have time to make it shorter. Your aim should be to write simple clear text that explains your thesis in a precise and elegant manner.

Formatting

It is essential to be consistent throughout your thesis. You should not use "Fig. 1" in one chapter, "Figure 1" in another and "figure 1" in yet another. Inconsistencies such as these are annoying and make the reader think you could have been equally sloppy in your mathematics, computer code and numerical results.

□ The format of my thesis is consistent from start to end.

The notion of consistent format must also be extended to mathematical notation. The thesis should present an integrated defense of a clear thesis statement; the mathematical notation should not change from one section to another based on the thoughtlessness of the writer or a different source. This applies to scalars and vectors, location, random variables, indexing and so on.

□ The mathematical notation and symbols are consistent from start to end.

References are important. It is unacceptable to present another's results or ideas as your own. You should not borrow text or figures or ideas without giving credit where credit is due. References to precise ideas/figures are straightforward. It is less clear how to reference a widely used technique. In general, reference the original source, an exceptionally important reference that popularized or explained the concept and a recent review text.

□ The original source of all specific and general concepts that are not mine are cited.

Thorough proof reading is important. Most professors read for content, but get distracted by grammatical and spelling errors. We rely too heavily on Word's spell check. Automatic corrections may not be correct. Many legitimate technical words and word usages are not acceptable to Word. Mistakes are often not highlighted, for example, the word 'from' typed as 'form'. A careful objective proof read is required.

□ I have thoroughly read my thesis looking for grammar and spelling errors.

Figures

The GSLIB PostScript generating code has been useful, but has also promoted poor practice in figure presentation. The text was designed to be readable at a nominal 4x3 inch format. Figures are often reduced for presentation. The reduced text is unreadable. GSLIB is not the only culprit. Many people shrink their figures to match the size of journal articles or newspapers. Everything should be readable – the text should be clear (nothing below 8 point font) and all lines and text should be clear. If the information is worth showing, then it should be readable.

□ All text and symbols on all figures are readable.

The best size for a figure does not change from the beginning of your thesis to the end. If a figure is worthy of one half page in one place, then it is worthy of a half page elsewhere. All figures showing similar content should be the same size.

□ The size of all figures of the same basic content are the same size.

All axes must be labeled. Schematic figures or flowcharts do not require a scale, but charts or maps with axes, points, or lines must be labeled. The reader should not have to read too carefully or employ detective skills to establish what is being plotted. The distance scale should be given somehow. It is always good practice to show a graphical scale that can withstand the figure being enlarged or reduced. The minimum and maximum coordinates should be labeled on all figures with spatial coordinates.

□ All axes, symbols and lines on all figures are labeled or are absolutely clear.

Often, figures from other sources are used or redrawn. If there is any question, then you should carefully cite the original source for all figures. This includes generic drawings off the internet or from your previous publications that have been published elsewhere. You sign the copyright to your own material over to a publisher when you publish a paper or book. Many people do not seek permission to use their own material, but care should be taken.

□ All external sources of information on my figures have been cited.

Disclaimer

You could check all of the boxes and your thesis may still require major changes. You could violate some rules and have a great thesis. That is why a group of professionals that have gone through the process and reviewed many other theses are judging yours. This document presents some guidelines and not a recipe.