MSc Research Skills

Lecture: From research proposal to MSc thesis

D G Rossiter

University of Twente.
Faculty of Geo-information Science & Earth Observation (ITC)

December 10, 2012

Outline

Topic: Proposal vs. thesis

Recall: The purpose of a research **proposal** is to:

- **convince** the research sponsor that you know the previous work on a subject;
- and that you have a **workable plan** on how to go beyond it and **contribute** to science via research.

But, the purpose of a research **thesis** is to:

- **report** on a **research project** . . .
- . . . according to defined **standards**.
- A thesis is a **scientific document**.

Structure of the research proposal vs. thesis

Recall: The **research proposal** has a conventional structure:

Problem ⇒ **Objectives** ⇒ **Questions** ⇒ **Hypotheses** ⇒ **Methods**

The **thesis** must then contain four more elements:

Results ⇒ **Discussion** ⇒ **Conclusions** ⇒ **Recommendations**

Note these are *not* necessarily section titles! Just conceptual headings.

These elements may be organized in several ways.

Topic: Thesis structure

The thesis is the **story** of a research project; so the **structure** of the document depends on the **most effective** way to **tell the story**.

Two structures are:

1. Straight-line
2. Parallel based on sub-topics

A simple structure: one main line

1. Introduction (problems, objectives, questions, hypotheses)
2. Literature review
3. Study area (if relevant); Data description (if relevant)
4. Methods
5. Results & Discussion
6. Conclusion & Recommendations

"Results" and "Discussion" may be in one Chapter or separated.

Structure based on sub-topics

Sometimes the research project can be naturally divided into a several **sub-topics** which follow in **sequence**.

This can be the basis of an effective thesis structure.

Example 1: Geothermal exploration

1. Introduction
2. Geothermal exploration – a review
3. Study area
4. Conceptual model of geothermal prospectivity
5. Analysis of geophysical data for indications of geothermal prospectivity
6. Analysis of Landsat TM data for indications of geothermal prospectivity
7. Regional-scale predictive modelling of geothermal prospectivity
8. Conclusions and recommendations

(source: Hendro Wibowo (2006): "Spatial Data Analysis and Integration for Regional-Scale Geothermal Prospectivity Mapping, West Java, Indonesia")

Example 2: Flood modelling

1. Introduction
2. Research procedure
3. Literature review
4. Study area
5. Soil properties in relation to land use (*note*: **plot** scale)
6. Surface runoff modelling (*note*: **hillslope** scale)
7. Flood modelling (*note*: **catchment** scale)
8. Flood hazard assessment with land use change scenarios (*note*: **integrates** scales)
9. Conclusions & Recommendations

(continued . . .)

Example continued

Note the **sequence of scales**; The results of modelling at finer scales are inputs to the model at the next coarser scale.

These are then followed by an **integrating** chapter: flood hazard (catchment) as affected by land use changes (plot), as revealed by the three-step modelling.

(source: Saowanee Prachansri (2007): "Analysis of soil and land cover parameters for flood hazard assessment : a case study of the Nam Chun watershed, Phetchabun, Thailand")

Topic: Results & Discussion

Prologue: recall Julius Caesar's report of one of his "research projects":

Veni, ***vidi***, ***vici***

- *Veni*: "I came": **Methods**
- *Vidi*: "I saw": **Results**
- *Vici*: "I conquered": **Discussion**

Results vs. Discussion

1. The **results** are what was actually observed when methods were applied
   - Results are presented **neutrally** (writing style is "reporting")
     For example:
     "Spectra of the tree species were almost identical in the winter image, but showed major differences in the summer image."
2. The **discussion** places these in scientific context.
   - Discussion is the reasoned **opinion** of the author (writing style is "argument").
     For example:
     "All species were dormant in the winter, so that characteristic absorption features, for example of chlorophyl, were absent."

Discussion points

Here are some questions that might lead to discussion:

- What is the **interpretation** of this result?
- If the result is presented as a **figure or table**, what is the reader supposed to infer from it?
- Is the result as **expected** (hypothesized)? If not, why not?
- Is the result in agreement with **previous research**? If not, why not? (What makes this case different?)

Discussing figures and tables

Figures and tables must be **referred to** in the text, and then **interpreted**, e.g.:

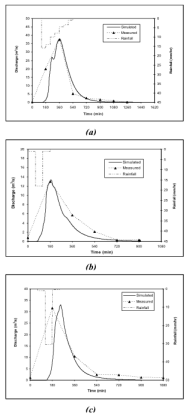


Figure 6-3: Measured and simulated discharge in Nam Chun catchment on 060905(a), 180905(b) and 260905(c) events.

Table 6-8: Observed and simulated peak discharge in Nam Chun catchment.

Events	Rainfall(mm)	Peak discharge	
		Obs(m^3/s)	Sim(m^3/s)
Calibration			
60905	52.58	37.5	37.90
180905	18.43	13.00	13.40
260905	29.49	31.45	32.99
Validation			
70905	16.60	4.2	4.72
90905	10.61	1.02	1.30
120905	26.74	15.37	23.81

"Model results were compared with measured discharge at the catchment outlet. The simulated and measured hydrographs for three events are shown in **Figure 6-3**, and their comparative summary statistics in **Table 6-8**. The model **closely fits the peak discharge** volume and time, **except for** the 26-Sept-2005 event, where the predicted peak is too large and early. **This is likely due to** imprecision in the measured hydrograph, due to the sparse recording interval (every three hours)." – adapted from Prachansri (2007), §6.3

Revisiting the literature review

Recall: the literature review includes **related work** which **justifies** the current research.

This work should also be **compared** to related work.

Because of the **time lapse** between proposal and thesis (four to six months), there may well be **new literature** relevant to the topic.

So before writing the discussion, **the literature search should be repeated** using the same search strategy as during the proposal stage

The **literature review should be updated** with the new references.

Organization of the results & discussion

Two possibilities:

1. **Sequential**: Results followed by Discussion
2. **Parallel**: each result with its own discussion.

Parallel structure

1. Results

 1.1 Result for question 1
 1.2 Result for question 2
 1.3 . . .
 1.n Result for question n

2. Discussion

 2.1 Discussion of result 1, with respect to question 1
 2.2 Discussion of result 2, with respect to question 2
 2.3 . . .
 2.n Discussion of result n, with respect to question n

Example of parallel structure

1. Methods
 1.1 Building footprint extraction
 1.2 Classification accuracy assessment
 1.3 Detection of new buildings

2. Results
 2.1 Building footprint extraction
 2.2 Classification accuracy assessment
 2.3 Detection of new buildings

3. Discussion
 3.1 Building footprint extraction
 3.2 Classification accuracy assessment
 3.3 Detection of new buildings

(source: Du Ye (2008) "Verification of tsunami reconstruction projects by object-oriented building extraction from high resolution satellite imagery")

Sequential structure

Results and discussion for each question are written together:

1. Results & Discussion

 1.1 Result for question 1; discussion with respect to question 1
 1.2 Result for question 2; discussion with respect to question 2
 1.3 . . .
 1.n Result for question nl discussion with respect to question n

Topic: Common mistakes in the results & discussion

- **Under**-interpretation
  - not getting full value from the results
- **Over**-interpretation
  - making unsubstantiated claims

Under-interpretation

- Results must be **interpreted**, not just **presented**.
- Every **table** and **figure** must be discussed.
  - this is *not* a repetition of the table or figure contents; the reader can see this for themselves
  - it is drawing attention to the **outstanding** (most important) results ...
  - ... and **explaining** them.

Example

Species	Slope	SE slope	R^2	n
Acacia erioloba	0.6133	0.0163	0.983	24
Acaia fleckii	0.6757	0.0238	0.971	24
Acacia luederitzii	0.615	0.0135	0.99	21
Burkea africana	0.7692	0.0091	0.998	18
Boscia albitrunca	0.7121	0.0268	0.967	24
Dichrostachyus cinerea	0.392	0.0077	0.991	23
Lonchocarpus nelsii	0.7943	0.0132	0.995	18
Ochna pulcra	0.6581	0.011	0.994	23
Terminalia sericea	0.5317	0.009	0.995	18

Table 1: Per-species linear models, Sapwood area vs. Stem area

"Table 1 shows the linear regression coefficients, their standard errors, coefficients of determination, and number of observations, for the prediction of cross-sectional sapwood area from cross-sectional stem area for the nine species."

Discussion of this table

Wrong: "The slope for *Acacia erioloba* was 0.6133 ± 0.0163, with $R^2 = 0.983$ (n=24), for *Acaia fleckii* 0.6757 ± 0.0238, with $R^2 = 0.971$ (n=24) ..."

Somewhat better: "Slopes ranged from 0.392 (*Dichrostachyus cinerea*) to 0.7692 (*Burkea africana*); standard errors of the slopes from ...; R^2 from ...

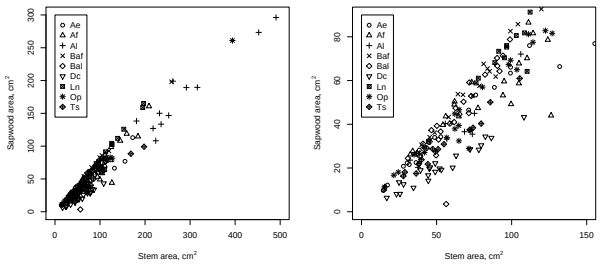
Better: "Slopes varied by a factor of almost two, from 0.392 (*Dichrostachyus cinerea*) to 0.7692 (*Burkea africana*); standard errors were all quite small (0.078 to 0.024) relative to the slopes. All models explained almost all the variance ($R^2 > 0.971$)."

Best: "Slopes varied by a factor of almost two, from 0.392 (*Dichrostachyus cinerea*) to 0.7692 (*Burkea africana*). This large variation is due to the major differences in tree morphology. *Dichrostachyus* species have very thick trunks relative to their height ... This result clearly shows that these relations must be species-specific."

Discussion of graphics

- **Draw reader's attention** to the outstanding features shown in the graphic
- **Interpret**:
  - what does this imply about **"nature"** (the thing being studied)?
  - what does this imply about the **analysis** (steps to be followed)?

Example



"Figure 1 shows the relation between sapwood area and stem area for the nine species. There is generally a linear relation, especially for the smaller trees; however for the largest trees there seems to be a smaller increase in sapwood for a corresponding increase in stem area. Further, the relation for *Dichrostachyus cinerea*) is clearly **anomalous** . . . These discrepencies **can be explained by** . . .

'It is clear that not all species have the same relation, even if we consider only the smaller trees. Thus, **per-species statistical relations must be developed**."

Over-interpretation

Statements must be supported by **your** results . . .

. . . possibly **in conjunction with** results from other studies.

Example:

"Nowadays coastal areas are affected by **increasing frequency of extreme events** like tsunami, storm surges and cyclones **as a result of global climate change**.

- Where is the proof of "increasing frequency"?
- If this is proven, where is the proof of "as a result of global climate change"?
- Both of these are very hard to prove or even suspect, given the short time-series.
- Are they required by **your** study?

Moderation in interpretation

- Stick to **facts** and **direct inferences** from these:

  "Major tsumanis have affected the . . . coast in 1865, 1920, 1985 and 2007 [reference]. The last-named resulted in . . . deaths and . . . Rp. damage [reference]. As the population in the coastal areas has steadily increased [reference], combined with the national policy on concentrating economic activity in these areas [reference], vulnerability to tsunamis has increased accordingly."

Facts vs. interpretations

Do not be afraid to interpret, but do not extrapolate beyond what the evidence suggests:

- **Statement of fact**: "The usability test with planning staff in Province X was successful: 80% of the participants (18 of 20) could complete the tasks well within the alloted time"
- **Reasonable interpretation**: "Since this province's planning department was selected as representative (see Methods, §2.2), we expect that similar results would be obtained in other provinces; therefore the system seems ready for country-wide implementation."
- **Excessive interpretation**: "Planning agencies in all Southeast Asia should immediately implement the planning support system developed during this thesis project".

Topic: Conclusions & recommendations

The most interesting section of the thesis for many readers is the **conclusion**.

What finally does the author conclude about their work?

Also, can the author make any **recommendations** about how better to address the research question, or what follow-up steps should be taken?

Two structures:

- **Combined**: because the recommendations flow directly from conclusions;
- **Separate**: conclusions about the **present work**; after that recommendations for **future** work.

Conclusions

The conclusions are a **summary** of the results and discussion, *without* any justification.

(Readers who want justification will look back into the body of the thesis.)

Conclusions refer to the **objectives** and **answer the questions** posed in the introduction.

Note: It may not be possible to answer all the questions fully; the reason for this unsatisfactory conclusion is presented in the discussion of the relevant question.

Some questions for the conclusions

- Were the research questions proper and sufficiently specific to be addressed?
- Were the methods applied satisfactory for the purpose of answering the research questions? If not, what should have been done instead?
- Were the data collected sufficient? If not, what additional data should have been collected?
- Was the case study or study area appropriate to answer the questions? If not, what characteristics should have been changed?
- How widely are these results applicable? I.e. how generic are they? If the same methods were applied to other cases, would similar results be expected? Why or why not?
- To what degree do the results answer the question? If not fully, what further information is required to do so?

Recommendations

After spending substantial time with a research topic, the author should have developed some ideas about what should be done next:

- In **further research**;
- In **practice**, based on this research.

Some questions that can lead to recommendations:

- Should any **action** be taken based on the results of this work?
  - For example, should a methodology developed in a research project be operationalized?
  - If so, what **modifications** might be needed, who should do this, etc.

(continued . . .)

- Does this research suggest **followups**?
  - "We have come this far, the next step is . . . "
- Were there **mistake** in planning (e.g. sampling strategy), methods applied, logistics . . . ?
- What should be done to overcome any **limitations** in the present work?
- Is the work complete, and the problem **solved**?
  - Then the recommendation is to move on to something else.
- If this or similar work should be **re-done**, what should be changed from the way you went about it?

SWOT approach to conclusions and recommendations

A useful conceptual framework is the so-called SWOT:

Strengths What did the research accomplish well?

Weaknesses What did the research not accomplish so well, or what were its limitations?

Opportunities What paths does this research open up for us?

Threats What other approaches could be better to address this problem?

Of course, these sections are not written with these headings, they are to help you think about your research in context.

Topic: "Selling" your thesis

The thesis must be **sound science**, but it should also **convince** the reader that:

- the work is **important**;
- the **proper questions** have been asked;
- **proper methods** have been applied;
- you have **properly interpretated** the results, leading to strong (but not exaggerated!) **conclusions**;
- your **recommendations** are supported by the research.

How to "sell"

- Clear **structure** (outline, paragraphs with topic sentences)
- Short and **to the point** (easy to read)
  - without sacrificing relevant detail
- Clear, concise **language**; sound **logic** and **argumentation**
  - No vague statements; use the right modal qualifiers
- Statements with the right **strength**, depending on the facts
- Clear relation with **related work** (in introduction, discussion, conclusion)
  - Why is your work important? What did the others not do?
  - How are your results related to others? Confirm? Modify? Reject? their work.
  - Who is "right"?

Final thoughts

- Proper reporting is difficult
- Good writing is difficult
- Saying just what should be said (no less, no more) is difficult
- Finding the right phrases and words is difficult
- **Take enough time** to write and **revise**
- Write your thesis as a **story of a research project**, aiming the story at your **expected reader**
  - Graduate scientist with a background in your field, but not necessarily an expert in your topic