# Automating Condition Monitoring of Vegetation on Railway Trackbeds and Embankments

Roger G. Nyberg

A thesis submitted in partial fulfilment of the requirements of Edinburgh Napier University for the award of Doctor of Philosophy

June 2015

I hereby declare that the work presented in this thesis was solely carried out by myself at Edinburgh Napier University, Edinburgh, except where acknowledgements are made and that it has not been submitted for any other degree.

Roger G. Nyberg

<table>
<tr><td colspan="2">Title: Automated Condition Monitoring of Vegetation on Railway Trackbeds and Embankments</td></tr>
<tr><td colspan="2">Author: Roger G. Nyberg Student Matriculation Number at Edinburgh Napier University: 10020460</td></tr>
<tr><td colspan="2">Director of Studies: Professor Narendra K. Gupta, e-mail: N.Gupta@napier.ac.uk</td></tr>
<tr><td>Supervisor: Prof. Mark Dougherty, Dalarna University, e-mail: mdo@du.se</td><td>Additional Supervisor: Assoc. Prof. Siril Yella, Dalarna University, e-mail: sye@du.se</td></tr>
</table>

## Abstract

Vegetation growing on railway trackbeds and embankments present potential problems. The presence of vegetation threatens the safety of personnel inspecting the railway infrastructure. In addition vegetation growth clogs the ballast and results in inadequate track drainage which in turn could lead to the collapse of the railway embankment.

Assessing vegetation within the realm of railway maintenance is mainly carried out manually by making visual inspections along the track. This is done either on-site or by watching videos recorded by maintenance vehicles mainly operated by the national railway administrative body.

A need for the automated detection and characterisation of vegetation on railways (a subset of vegetation control/management) has been identified in collaboration with local railway maintenance subcontractors and Trafikverket, the Swedish Transport Administration (STA). The latter is responsible for long-term planning of the transport system for all types of traffic, as well as for the building, operation and maintenance of public roads and railways. The purpose of this research project was to investigate how vegetation can be measured and quantified by human raters and how machine vision can automate the same process.

Data were acquired at railway trackbeds and embankments during field measurement experiments. All field data (such as images) in this thesis

work was acquired on operational, lightly trafficked railway tracks, mostly trafficked by goods trains. Data were also generated by letting (human) raters conduct visual estimates of plant cover and/or count the number of plants, either on-site or in-house by making visual estimates of the images acquired from the field experiments. Later, the degree of reliability of (human) raters' visual estimates were investigated and compared against machine vision algorithms.

The overall results of the investigations involving human raters showed inconsistency in their estimates, and are therefore unreliable. As a result of the exploration of machine vision, computational methods and algorithms enabling automatic detection and characterisation of vegetation along railways were developed. The results achieved in the current work have shown that the use of image data for detecting vegetation is indeed possible and that such results could form the base for decisions regarding vegetation control. The performance of the machine vision algorithm which quantifies the vegetation cover was able to process 98% of the image data. Investigations of classifying plants from images were conducted in in order to recognise the specie. The classification rate accuracy was 95%.

Objective measurements such as the ones proposed in thesis offers easy access to the measurements to all the involved parties and makes the subcontracting process easier i.e., both the subcontractors and the national railway administration are given the same reference framework concerning vegetation before signing a contract, which can then be crosschecked post maintenance.

A very important issue which comes with an increasing ability to recognise species is the maintenance of biological diversity. Biological diversity along the trackbeds and embankments can be mapped, and maintained, through better and robust monitoring procedures. Continuously monitoring the state of vegetation along railways is highly recommended in order to identify a need for maintenance actions, and in addition to keep track of biodiversity. The computational methods or algorithms developed form the foundation of an automatic inspection system capable of objectively supporting manual inspections, or replacing manual inspections.

# List of Publications

During this study, a number of scientific articles, conference proceedings and reports were published. These are listed below:

- Nyberg, R. G., Gupta, N. K., Yella, S., Dougherty, M. S. *Machine vision for condition monitoring vegetation on railway embankments*, Proceedings of the 6th IET Conference on Railway Condition Monitoring RCM 2014. 17-18 Sept. Birmingham, U.K. Inproceedings/Conference paper
- Cederlund, H, Fogelberg, F., Hansson, D., Nyberg, R. G., Schroeder, H. *Development of Methods to Support Decisions for Weed Control on Railway Embankments*, 2014. Technical Report.
- Nyberg, R. G., Gupta, N. K., Yella, S., Dougherty, M. S. *Monitoring Vegetation on Railway Embankments : Supporting Maintenance Decisions*, Proceedings of the 2013 International Conference on Ecology and Transportation (ICOET 2013), Scottsdale, Arizona, USA. Conference paper
- Yella, S., Nyberg, R. G., Payvar, B., Dougherty, M., Gupta, N. K. *Machine vision approach for automating vegetation detection on railway tracks*, Journal of Intelligent Systems, 2013, Vol. 22, No. 2, 179-196. Article.
- Nyberg, R. G., Gupta, N., Yella, S., Dougherty, M. *Detecting Plants on Railway Embankments*, Journal of Software Engineering and Applications, 2013, Vol. 6, No. 3B, 8-12. Article

# Contents

Terminology, Definitions, and Symbols . . . . . . . . . . . . . . . 22

**I Introduction** **31**

**1 Introduction** **32**

1.1 Problem Description . . . . . . . . . . . . . . . . . . . . . . 34

1.1.1 Procurement of Railway Maintenance . . . . . . . . 40

1.2 Manual Inspection Routines . . . . . . . . . . . . . . . . . . 43

1.2.1 Administrative Regulations . . . . . . . . . . . . . . 44

1.2.2 Estimate of Inspection Costs . . . . . . . . . . . . . 45

1.2.3 Disadvantages with Manual Inspections . . . . . . . 45

1.3 Proposed Automated Solution . . . . . . . . . . . . . . . . 47

1.4 Research Questions . . . . . . . . . . . . . . . . . . . . . . 48

1.5 Aims and Objectives . . . . . . . . . . . . . . . . . . . . . . 49

1.5.1 Aims . . . . . . . . . . . . . . . . . . . . . . . . . . 49

1.5.2 Objectives . . . . . . . . . . . . . . . . . . . . . . . 50

1.5.3 Effects . . . . . . . . . . . . . . . . . . . . . . . . . 50

1.6 Research Methodology and Design . . . . . . . . . . . . . . 51

1.6.1 General Methodological Approach . . . . . . . . . . 51

1.6.2 Research Design . . . . . . . . . . . . . . . . . . . 55

1.7 Scope and Limitations . . . . . . . . . . . . . . . . . . . . . . 57
1.7.1 Scope . . . . . . . . . . . . . . . . . . . . . . . . . 57
1.7.2 Limitations . . . . . . . . . . . . . . . . . . . . . . 58
1.8 Outline of the Thesis . . . . . . . . . . . . . . . . . . . . . . 58

**II Technology and Literature Review 60**

**2 Maintenance Strategies 62**
2.1 Vegetation Monitoring . . . . . . . . . . . . . . . . . . . . . . 65
2.2 Sensor Technology . . . . . . . . . . . . . . . . . . . . . . . 66
2.2.1 Sensor Technology for Detecting Vegetation . . . . . 67
2.3 Conclusions & Discussion . . . . . . . . . . . . . . . . . . 70

**3 Plant Science 71**
3.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . 71
3.2 Plant Requirements . . . . . . . . . . . . . . . . . . . . . . 71
3.3 Analysing and Classifying Plants . . . . . . . . . . . . . . . 72
3.4 Biophysical Data, and Root and Shoot Ratios . . . . . . . . 73
3.5 Common Woody Plants along Railways . . . . . . . . . . . 75
3.5.1 Woody Plant Candidates for Data Collection . . . . 77

**4 Measuring Terrestrial Vegetation 79**
4.1 Sampling Units . . . . . . . . . . . . . . . . . . . . . . . . . 79
4.2 Spatial Plant Distribution . . . . . . . . . . . . . . . . . . . 80
4.3 Vegetation Attributes . . . . . . . . . . . . . . . . . . . . . . 83
4.3.1 Plant Cover . . . . . . . . . . . . . . . . . . . . . . . 84
4.3.2 Plant Frequency . . . . . . . . . . . . . . . . . . . . 90
4.3.3 Plant Density . . . . . . . . . . . . . . . . . . . . . . 92
4.3.4 Plant Biomass: A Measure of Production . . . . . . 93
4.4 Conclusions & Discussion . . . . . . . . . . . . . . . . . . 93

# III Data Collection 95

## 5 Field Measurements and Experiments 96

5.1 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
5.1.1 Camera Sensors . . . . . . . . . . . . . . . . . . . . . 97
5.1.2 Information Storage . . . . . . . . . . . . . . . . . . . 98
5.2 Motivation for the Field Experiments . . . . . . . . . . . . . 98
5.3 $1^{st}$ Field Experimental Site: Falun - Grycksbo, Sweden . . . 102
5.3.1 Motivation for Conducting the 1st Field Experiment . 103
5.3.2 Method . . . . . . . . . . . . . . . . . . . . . . . . . 104
5.3.3 Results and Conclusions of the 1st Field Experiment 108
5.4 $2^{nd}$Field Experimental Site: Falun - Grycksbo, Sweden . . . 110
5.4.1 Results and Conclusions of the 2nd Field Experiment 112
5.5 $3^{rd}$ Field Experimental Site: Älvdalsbanan, Oxberg . . . . . 113
5.5.1 Results and Conclusions of the 3rd Field Experiment 115
5.6 $4^{th}$ Field Experimental Site: Vetlanda . . . . . . . . . . . . 115
5.7 $5^{th}$ Field Experiment . . . . . . . . . . . . . . . . . . . . . 118
5.7.1 Results and Conclusions of the 4th Field Experiment 119
5.8 Indoor Laboratory Experiments . . . . . . . . . . . . . . . . 121
5.8.1 Motivation for Laboratory Experiments . . . . . . . . 121
5.9 Conclusion and Discussion . . . . . . . . . . . . . . . . . . 124

# IV Manual Assessments of Terrestrial Vegetation 126

## 6 Visual Estimates, Reliability & Raters Agreement 128

6.1 Recording Human Estimates . . . . . . . . . . . . . . . . . 128
6.2 Levels of Measurement . . . . . . . . . . . . . . . . . . . . 129
6.2.1 Data Transformation . . . . . . . . . . . . . . . . . . 133

6.3 Raters' Agreement/Reliability . . . . . . . . . . . . . . . . . . 136
6.4 Counting Plant Clusters/Patches . . . . . . . . . . . . . . . . 140
6.5 Six Studies in Reliability of Visual Estimates . . . . . . . . . . 142
6.6 1st Study of Visual Estimates of Images . . . . . . . . . . . 142
6.6.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 142
6.6.2 Results and Conclusions of the 1st Study . . . . . . 143
6.7 2nd Study of Visual Estimates On-site . . . . . . . . . . . . . 145
6.7.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 145
6.7.2 Results and Conclusions of the 2nd Study . . . . . . 146
6.8 3rd Study of Visual Estimates of Images . . . . . . . . . . . . 148
6.8.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 148
6.8.2 Visual Estimates of Plant Cover and Sub-plot Frequency . . . . . . . . . . . . . . . . . . . . . . . . 150
6.8.3 Counting Plant Clusters . . . . . . . . . . . . . . . . 164
6.8.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . 168
6.9 4th Study of Visual Estimates of Images . . . . . . . . . . . 170
6.9.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 170
6.9.2 Visual Estimates of Plant Cover . . . . . . . . . . . . 172
6.9.3 Counting Plant Clusters . . . . . . . . . . . . . . . . 181
6.9.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . 186
6.10 5th Study of Visual Assessments: Älvdalsbanan, Oxberg . . 189
6.10.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 189
6.10.2 Results and Conclusions of the 5th Study . . . . . . 190
6.11 6th Study of Visual Estimates On-site in Vetlanda . . . . . . 192
6.11.1 Method . . . . . . . . . . . . . . . . . . . . . . . . 192
6.11.2 Results and Conclusions of the 6th Study . . . . . . 193
6.12 Conclusion and Discussions . . . . . . . . . . . . . . . . . . 194

**V Machine Vision and Machine Learning 195**

**7 Classification in Machine Learning for Machine Vision 197**

7.1 Recognising Objects and the Classification Model . . . . . 198

7.1.1 Classification using Bag-of-Features . . . . . . . . . 201

7.2 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . 203

**8 Quantification of Vegetation using Machine Vision 204**

8.1 Quantification of Vegetation using Machine Vision . . . . . 204

8.1.1 Algorithm Description . . . . . . . . . . . . . . . . . 205

8.1.2 Machine Vision Algorithm Results . . . . . . . . . . 209

8.1.3 Comparison: MV Algorithm vs. Manual Visual Estimates . . . . . . . . . . . . . . . . . . . . . . . . 210

8.1.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . 212

**9 Estimating the Quantity of Biomass along Railway Sections 214**

9.1 Correlation between Root and Shoot Weights and Plant Cover 214

9.2 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . 217

9.3 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . 221

**10 Recognition and Characterisation of Woody Plants using Machine Vision 223**

10.1 Method . . . . . . . . . . . . . . . . . . . . . . . . . . . 224

10.2 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . 225

10.3 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . 227

**VI Conclusions and Discussion 228**

**11 Conclusions and Discussion 229**

**12 Future Work** **236**

**References** **241**

**A Data from Falun - Grycksbo Railway, Sweden** **254**

A.1 Plant Root and Shoot Dry Weights . . . . . . . . . . . . . . 254

**B Data from Älvdalsbanan, Oxberg, Sweden** **256**

B.1 Plant Root and Shoot Dry Weights . . . . . . . . . . . . . . 256

**C Data from Studies of Reliability of Visual Estimates** **262**

C.1 Data from the study made in section 6.6 . . . . . . . . . . . 262

C.2 Data from the study made in section 6.7 . . . . . . . . . . . 264

C.3 Data from the study made in section 6.8 . . . . . . . . . . . 267

C.4 Data from the study made in section 6.9 . . . . . . . . . . . 279

**D Botanic Taxonomic Hierarchy** **288**

**E Experimental Sites: Field Layer Classes and Tree Types** **289**

E.1 General description of experimental locations . . . . . . . . 289

**F Database Modelling** **291**

F.1 Database Model: Acquired Data Storage . . . . . . . . . . 291

**G Assistance when Estimating Plant Cover** **293**

**H Differences in Counting Plants Clusters** **294**

**I The Engineering Design Process & The Scientific Method** **296**

**J Flow-chart for estimating the negative binomial exponent $k$** **297**

J.1 Methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 297

**K Internet-based Sources** **300**

# List of Figures

1.1 Growing vegetation on the trackbed a) Vetlanda, Sweden, and b) Falun, Sweden . . . . . . . . . . . . . . . . . . . . . . 33
1.2 Overview of the track area . . . . . . . . . . . . . . . . . . . . . . 34
1.3 Vegetation growing on the trackbeds: a) Borlange, Sweden, and b) Mora, Sweden . . . . . . . . . . . . . . . . . . . . . . 35
1.4 Collapse of railway embankment . . . . . . . . . . . . . . . . 37
1.5 Leaf buster train (Network Rail UK) . . . . . . . . . . . . . . 38
1.6 Maintenance procurement process . . . . . . . . . . . . . . 41
1.7 General Research Process (Modified and redrawn from Oates (2006)) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52

2.1 Maintenance - Overall view (Redrawn from BSI (2010)) . . 63
2.2 Steps in condition-based maintenance . . . . . . . . . . . . 64
2.3 Output from a Yara N-sensor (Cederlund et al., 2014, pp. 48) 68

3.1 Picea abies seedling established on sleepers . . . . . . . . 76
3.2 Pinus sylvestris seedling established on a sleeper . . . . . 76
3.3 Betula seedling establishment on ballast . . . . . . . . . . . 77

4.1 a) Random, and b) aggregated, and c) uniform plant patterns 80
4.2 Example of distribution dependent on scale of sample square 83
4.3 Aerial cover vs. Basal cover Elzinga et al. (1998, p.178) . . 85

4.4 a) Aerial foliar cover vs. b) Aerial canopy cover Coulloudon et al. (1999, p.25) . . . . . . . . . . . . . . . . . . . . . . . 85
4.5 Three definitions of cover . . . . . . . . . . . . . . . . . . . 86
4.6 10-pin point frame (Bonham, 1989) . . . . . . . . . . . . . 90
4.7 Estimating frequency . . . . . . . . . . . . . . . . . . . . . . 92

5.1 Dominant field layer class for a certain area (Reprinted with permission) . . . . . . . . . . . . . . . . . . . . . . . . . . 100
5.2 Camera's vertical setup . . . . . . . . . . . . . . . . . . . . 102
5.3 Overview Grycksbo: Railway section under investigation . . 103
5.4 a) Sample plot area, and b) point quadrat grid frame containing 100 sub-plots . . . . . . . . . . . . . . . . . . . . . . 105
5.5 Subset of a sub-plot grid, or point quadrat grid . . . . . . . 107
5.6 Overview of the railway section . . . . . . . . . . . . . . . 111
5.7 Sample plot area at 2nd field experimental site. . . . . . . . 112
5.8 Overview of the sampled railway section (Oxberg, Sweden) 113
5.9 Sample plot no. 200 a) before, and b) after excavation . . . 114
5.10 Conifer woody plants, in plot 202 (Oxberg, Sweden) . . . . 114
5.11 Documentation of excavated plants in lab (Picea abies) . . 115
5.12 First part of the Vetlanda railway section . . . . . . . . . . . 116
5.13 Trolley with mounted camera . . . . . . . . . . . . . . . . . 117
5.14 Vetlanda sample plots a) ID 4, and b) ID 22 . . . . . . . . . 117
5.15 Sleeper spacing . . . . . . . . . . . . . . . . . . . . . . . . 118
5.16 Data - Sleeper spacings . . . . . . . . . . . . . . . . . . . . 120
5.17 The indoor lab setup . . . . . . . . . . . . . . . . . . . . . . 123
5.18 Examples of conifer plants and leafs on the object table . . 123

6.1 Boxplot . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133
6.2 a) Skewness, and b) kurtosis . . . . . . . . . . . . . . . . . 134

6.3 a) Ellipse, and b) Usage on plants/patches . . . . . . . . . 141
6.4 Interpretation of the definition of patches . . . . . . . . . . . 141
6.5 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 144
6.6 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 144
6.7 Randomised plots in a stratum . . . . . . . . . . . . . . . . 146
6.8 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 147
6.9 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 147
6.10 Visual aid chart for estimating cover . . . . . . . . . . . . . 150
6.11 Sample area with a 10 x 10 sub-plot frame . . . . . . . . . 150
6.12 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 151
6.13 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 152
6.14 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 153
6.15 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 153
6.16 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 154
6.17 VE per sample plot: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . 155
6.18 VE by each rater: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 156
6.19 VE per sample plot: a) original data b) log10-transformed data156
6.20 VE by each rater: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . 157

6.21 VE per sample plot: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . 158
6.22 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 159
6.23 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 159
6.24 VE by each rater: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 160
6.25 VE per sample plot: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . 161
6.26 VE by each rater: a) original data b) log10-transformed data 162
6.27 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 162
6.28 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 163
6.29 VE by each rater: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 164
6.30 Each rater's counting of plant clusters after the ACC session: a) original data, and b) log10-transformed data . . . . 165
6.31 Plant cluster counts per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . 165
6.32 Each rater's counting of plant clusters after the AFC session: a) original data, and b) log10-transformed data . . . . 166
6.33 Plant cluster counts per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . 167
6.34 a) Sampling area, and b) AFC vs ACC method . . . . . . . 171
6.35 VE by each rater: a) original data, and b) log10- transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 173
6.36 VE per sample plot: a) original data b) log10-transformed data173
6.37 VE by each rater: a) original data b) log10-transformed data 174
6.38 VE per sample plot: a) original data b) log10-transformed data175

6.39 VE by each rater: a) original data b) log10-transformed data 176
6.40 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 176
6.41 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 177
6.42 VE per sample plot: a) original data b) log10-transformed data 178
6.43 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 179
6.44 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 179
6.45 VE by each rater: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 180
6.46 VE per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . . . . . . 181
6.47 Each raters counting of plant clusters after the ACC session: a) original data, and b) log10-transformed data . . . . . . . 182
6.48 Plant cluster counts per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . 183
6.49 Each raters counting of plant clusters after the AFC session: a) original data, and b) log10-transformed data . . . . . . . 183
6.50 Plant cluster counts per sample plot: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . 184

7.1 A fictive class of locomotives . . . . . . . . . . . . . . . . . 202

8.1 View over the trackbed . . . . . . . . . . . . . . . . . . . . . 205
8.2 a) After Mean-Shift Clustering, and b) HOG mask on top of the decorrelated stretch result . . . . . . . . . . . . . . . . . 206
8.3 a) After Otsu thresholding, and b) Extrapolated skeletonised line segments representing the rails . . . . . . . . . . . . . 207
8.4 ROI mask layer placed on top of original image . . . . . . . 208
8.5 a) Perspective corrected image, and b) Extracted vegetation 209

8.6 Cover estimates in a) June, and b) August . . . . . . . . . . 211

8.7 Correlation manual VE vs. MV in: a) June, and b) August . 212

9.1 Overview, plot 200-204 (Oxberg, Sweden) . . . . . . . . . . 215

9.2 Histogram and density plots of a) root dry weights and b) shoot dry weights from sample plots 1 to 5 (Grycksbo, Sweden) 216

9.3 Histogram and density plots of a) root dry weights, and b) shoot dry weights from sample plots 200 to 204 (Oxberg, Sweden) . . . . . . . . . . . . . . . . . . . . . . . . . . . 217

9.4 Root dry weight vs. shoot dry weight data (Grycksbo, Sweden) 218

9.5 Root dry weight vs. shoot dry weight data (Oxberg, Sweden) 219

9.6 Total shoot dry weight (from 197 plants) vs. MV computed cover for the plots 200 to 204 (Oxberg, Sweden) . . . . . . 221

10.1 Detected image features on: a) Norway Spruce, b) Silver Birch leaf, c) negative image, and d) Alder leaf. . . . . . . . 225

11.1 Systematic sampling along the tracks . . . . . . . . . . . . 231

12.1 Proposed segments of the railway embankment . . . . . . . 237

B.1 The five sample plot areas . . . . . . . . . . . . . . . . . . . 257

C.1 a) Histogram and density plots of: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . 263

C.2 Density plot of the residuals from the log10-transformed data 264

C.3 Histogram and density plots of: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . 266

C.4 Density plot of the residuals from the log10-transformed data 266

C.5 a) Histogram and density plots of: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . 267

C.6 Density plot of the residuals from the log10-transformed data 268

C.7 a) Histogram and density plots of a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 268
C.8 Density plot of the residuals from the log10-transformed data 269
C.9 a) Histogram and density plots of a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . 270
C.10 Density plot of the residuals from the log10-transformed data 271
C.11 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 271
C.12 Density plot of the residuals from the log10-transformed data 272
C.13 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 273
C.14 Density plot of the residuals from the log10-transformed data 273
C.15 Histogram and density plots of: a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . 274
C.16 Density plot of the residuals from the log10-transformed data 275
C.17 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 275
C.18 Density plot of the residuals from the log10-transformed data 276
C.19 Histogram and density plots of a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . . . . 277
C.20 Density plot of the residuals from the log10-transformed data 277
C.21 Histogram and density plots of a) original data, and b) log10-transformed data . . . . . . . . . . . . . . . . . . . . . . . 278
C.22 Density plot of the residuals from the log10-transformed data 279
C.23 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 280
C.24 Density plot of the residuals from the log10-transformed data 280
C.25 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . 281
C.26 Density plot of the residuals from the log10-transformed data 282

C.27 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . . . 282
C.28 Density plot of the residuals from the log10-transformed data 283
C.29 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . . . 284
C.30 Density plot of the residuals from the log10-transformed data 285
C.31 a) Histogram and density plots of a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . . . 285
C.32 Density plot of the residuals from the log10-transformed data 286
C.33 Histogram and density plots of: a) original data, and b) log10-transformed . . . . . . . . . . . . . . . . . . . . . . . . 287
C.34 Density plot of the residuals from the log10-transformed data 287

F.1 Database model . . . . . . . . . . . . . . . . . . . . . . . . 292

G.1 Aiding template for the estimate of plant cover . . . . . . . 293

H.1 Differences in counting plant clusters between rater A to E . 295

J.1 Estimating the negative binomial exponent $k$, with reference to method ♯1 to ♯4 in section J.1. . . . . . . . . . . . . . . . 298

# List of Tables

3.1 A method of analysing plants by their physiognomy and structure devised by Kuchler (1966) . . . . . . . . . . . . . . . . 73
3.2 Plants likely to be found on railway embankments . . . . . . 78

4.1 Plant cover class systems (Mueller-Dombois and Ellenberg, 1974) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 88

5.1 Field Layer Types (Hagglund and Lundmark, 2004) . . . . 99
5.2 Suggested vegetation inventory classes by Below, Deutsche BahnHuisman (2001) . . . . . . . . . . . . . . . . . . . . . . 104
5.3 Estimating cover (%) using a) AFC, and b) sub-plot frequency 106
5.4 Non-guided visual cover estimation . . . . . . . . . . . . . . 109
5.5 VE cover estimates using a sub-plot grid quadrat . . . . . . 109
5.6 a) Spacings between sleepers. b) Sleepers widths . . . . . 119
5.7 Tree frequency in Sweden . . . . . . . . . . . . . . . . . . . 122
5.8 Measured average incoming light at the object . . . . . . . 123

6.1 Levels of measurement, (Siegel, 1957) . . . . . . . . . . . 132
6.2 The levels of agreement among the raters . . . . . . . . . . 145
6.3 The three raters' levels of agreement . . . . . . . . . . . . . 148
6.4 Overview - number of observations per person . . . . . . . 149
6.5 ICC(2,1) for agreement between the three raters' counting . 167

6.6 The raters' two (paired) sessions of counting plant clusters, df = 34 . . . 168
6.7 The three raters' level of agreement as of the ICC(2,1) . . . 169
6.8 Overview - number of observations per person . . . 171
6.9 The five raters' levels of counting agreement as of the ICC(2,1) 184
6.10 The raters' two (paired) sessions of counting plant clusters, df=50 . . . 186
6.11 Summary of difference in between raters from section 6.9.2 to 6.9.2 . . . 187
6.12 The five raters' level of agreement as of the ICC(2,1) . . . 187
6.13 Oxberg: on-site visual cover examination . . . 190
6.14 Difference in cover estimates per plot . . . 191
6.15 Agreement between the two domain raters' VE . . . 191
6.16 Raters average cover estimates . . . 193

8.1 Computed cover (%) by the algorithm along the entire railway section . . . 210

9.1 Computed cover (%) in sample plot 200 to 204 (Oxberg, Sweden) . . . 220

10.1 The categories to be classified . . . 224
10.2 BoF classification results . . . 226
10.3 Confusion matrix (%) for the SVM Chi$^2$kernel with DSIFT feature extraction. . . . 227

A.1 Root and shoot dry weights (in grams) collected at plot 1 to 5 (Grycksbo, Sweden) . . . 255

B.1 Root and shoot dry weights (in grams) collected from plots 200 to 201 (Oxberg, Sweden) . . . 258
B.2 Root and shoot dry weights (in grams) collected from plots 201 to 203 (Oxberg, Sweden) . . . 259

B.3 Root and shoot dry weights (in grams) collected from plots 203 to 204 (Oxberg, Sweden) . . . . . . . . . . . . . . . . . . 260
B.4 Root and shoot dry weights (in grams) collected from plot 204, continued. (Oxberg, Sweden) . . . . . . . . . . . . . . 261

C.1 The 13 raters' (A-M) cover estimates (%) from 9 images, I1-I9 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 263
C.2 The three raters (A to C) cover estimates (%) from 10 plots, I1 to I10 . . . . . . . . . . . . . . . . . . . . . . . . . . 265

D.1 Botanic taxonomic hierarchy from Mauseth (1998) . . . . . 288

E.1 Frequencies of field layer (%) . . . . . . . . . . . . . . . . . . . 290
E.2 Frequencies of tree types (%) . . . . . . . . . . . . . . . . . . . 290

I.1 The engineering design process and the scientific method 296

K.1 Internet-based sources in the literature review . . . . . . . . 300

# Terminology, Definitions, and Symbols

| | |
|---|---|
| Aphyllous | Having no leaves |
| Assessor | A person who provides a rating or assessment of some phenonomen. Also denoted as rater or observer. |
| Ballast | A bed of crushed stones used to drain water from the embankent, and to distribute the load from sleepers |
| Biomass | In this context the plant biomass, phytomass, is considered: Total weight of plant biomass |
| Biometry | The science of measuring and statistically analysing biological data |
| Blanket | A layer, or several layers, laid over the subgrade to give the trackbed its desired performance characteristics. The layers can include layers of granular material and geosynthetics. (Rail Safety and Standards Board Ltd, UK) |
| BoF | Bag-of-Features |
| BoW | Bag-of-Words |
| BVH | Banverket Handbok - Swedish Rail Administration Manual. Banverket was superseded by Trafikverket in 2010 (i.e. the Swedish Transport Administration) |
| CBM | Condition Based Maintenance, or Condition Based Monitoring |

Convex hull — The convex hull of a set X of points in the Euclidean plane is the smallest convex set that contains X. The convex hull may be visualized as the shape enclosed by a rubber band stretched around X.

Corrective Maintenance (CM) — Unscheduled maintenance or repair actions, performed as a result of failures or deficiencies, to restore items to a specific condition. Maintenance done to bring an asset back to its standard functional performance: Any maintenance activity which is required to correct a failure that has occurred or is in the process of occurring. This activity may consist of repair, restoration, or replacement of components and can typically be planned, estimated and scheduled proactively.
(Peters, R. W., Maintenance Benchmarking and Best Practices. Blacklick, OH, USA: McGraw-Hill Professional Publishing, 2006. p 520)

Deciduous — Annual or seasonal loss of all leaves from a tree or shrub; it is the opposite of evergreen. (World Encyclopedia. Philip's, 2008. Oxford Reference Online. Oxford University Press. Hogskolan Dalarna. 6 September 2011)

df — degrees of freedom. A term used in statistics to describe the number of independent comparisons that can be made between the variables in a study. (oxfordreference.com, 2012-11-22)

Embankment — Artificial mound of imported material generally made of selected earth, gravel, or stone. (California High-Speed Rail Authority). A bank of earth or stone built to carry a road or railway over an area of low ground. (Oxford Dictionary of English, 2011)

Epiphytes — Grows on another plants. Air plants, e.g Orchids

Forbs — Herbaceous flowering plant other than a grass (Source: Oxford Dictionary of English. Edited by Angus Stevenson. Oxford University Press, 2010. Oxford Reference Online. 5 Sept. 2011 )

Graminoids Grasses and grasslike plants

ICC See Intraclass-correlation coefficient

Inter-observer agreement Also called inter-rater reliability, or inter-rater agreement is the degree of agreement among several observers which are assessing some phenonomen. It gives a score of how much the observers agree in the ratings given by the observers. Inter-observer, or inter-rater implies “between raters”

inter-rater reliability See inter-observer reliability

Intraclass Correlation Coefficient (ICC) The ICC is a measure of the reliability of measurements or ratings.

Knowledge Short desc: Knowledge refers to the stored information or models used by a person or a machine to interpret, predict, and appropriately respond to the out side world. (Fischler, 1987)

laypersons A person without professional or specialized knowledge in a particular subject (Oxford Dictionaries, 2012-10-10)

LED Light Emitting Diodes

Md The statistical median

measurement The process of assigning symbols to observations in some consistent manner (S. Siegel,1988)

monitoring Observe and check the progress or quality of (something) over a period of time; keep under systematic review: equipment was installed to monitor air quality.
- maintain regular surveillance over: he was a man of routine and it was easy for an enemy to monitor his movements. Source: Oxford Dictionary of English. Edited by

Angus Stevenson. Oxford University Press, 2010. Oxford Reference Online. Hogskolan Dalarna. 5 Feb. 2011.

NVC — The National Vegetation Classification is a detailed phytosociological classification, which assesses the full suite of vascular plant, bryophyte and macro-lichen species within a certain vegetation type. It is based on about 35,000 samples of vegetation. Source: http://jncc.defra.gov.uk/page-4262 (Retrieved 27 July 2011)

Observer — A person who provides a rating or assessment of some phenonomen. Also denoted as rater or assessor.

Pathogens — Disease-causing agents, e.g. special kind of micro-organisms, viruses, rats etc etc

Pattern recognition — Pattern recognition is the act of of acquiring raw data and taking actions based on a particular the class (or category) of that pattern. Three interrelated fields often used in pattern recognition regression, interpolation and density estimation. (Duda et al., 2000)

PFI — See Point frame interception

Plant Frequency — The proportion of sample areas in which the target specie is present. For each sample area it is a boolean decision, either the target specie is present, or absent. The frequency is the sum of all sample areas having the target specie percent divided by all the examined sample areas.

Point Frame Interception (PFI) — Method used for estimating abundance.

Point Quadrat Interception (PQI) — Method used for estimating abundance. Each time an individual specie is observed under an intersection in the quadrat grid (i.e. a cross-hair) it is counted as a hit, otherwise miss.

PQI — See Point quadrat interception

Predictor variable A predictor variable is a variable used in regression to predict another variable. It is sometimes referred to as an independent variable if it is manipulated rather than just measured. (http://onlinestatbook.com, 2014-11-30)

Preventive maintenance (PM) Maintenance carried out at predetermined intervals, or to other prescribed criteria, and intended to reduce the likelihood of a functional failure. Actions performed in an attempt to keep an item in a specific operating condition by means of systematic inspection, detection, and prevention of incipient failure; an equipment maintenance strategy based on replacing, overhauling, or remanufacturing an item at a fixed interval, regardless of its condition at the time. Scheduled restoration tasks and scheduled discard tasks are both examples of preventive maintenance tasks. See also scheduled maintenance. [Source: Peters, Ralph W. (Author). Maintenance Benchmarking and Best Practices. Blacklick, OH, USA: McGraw-Hill Professional Publishing, 2006. p 534.]

Quadrat A basic sampling unit of vegetation surveys. The sampling frame can be of any shape which not necessarily is a geometric regular quadrilateral. It can be rectangular, or circular as well.

Rater A person who provides a rating or assessment of some phenonomen. Also denoted as assessor or observer.

Reliability Is, generally, the proportion of real information about a construct of interest captured by measurements of it. E.g., if someone reported the reliability of their measure was 0.8, you could conclude that 0.8 of the variability in the scores captured by that measure represented the construct, and 0.2 represented random variation. The more uniform your measurement, the higher reliability will be. ( http://neoacademic.com , Retrieved 2 Feb 2014)

Residual The difference between a data observation and its corresponding fitted value obtained by regression analysis.

| | |
|---|---|
| ROI | Region of Interest. Selected samples of an image, such as areas containing vegetation, that are identified for a particular purpose |
| RQ | Research question |
| RSSB | Rail Safety and Standards Board (UK) |
| Single Point Interception (SPI) | Method used for estimating abundance. Each sample point is defined by a sampling pole positioned vertically to the ground |
| SNFI | Swedish National Forest Inventory |
| STA | The Swedish Transport Administration, Trafikverket, is responsible for long-term planning of the transport system for all types of traffic, as well as for building, operating and maintaining public roads and railways. |
| Subgrade | The upper part of the earthworks or natural ground on which the blanket layer rests. The subgrade includes any capping layer (prepared subgrade) designed to alter the stiffness of the subgrade. (Rail Safety and Standards Board Ltd, UK) |
| SYNOP | Report of surface observation from a land station according to World Meteorological Organization standards. |
| Trackbed | A general term referring to the ballast, blanket and subgrade. (Rail Safety and Standards Board Ltd, UK) |
| Trafikverket | Swedish for the Swedish Transport Administration, see STA. |
| Transect | A sample area usually in the form of a long continuous line in some direction. Along this line one counts and records occurrences of e.g. plants. |

| | |
|---|---|
| VE | Visual Estimation |
| vermin | Animal species which are regarded as pests or by man considered as troublesome and/or annoying. Especially associated to species able of carrying disease |
| VINNOVA | VINNOVA is a Swedish government agency working under the Ministry of Enterprise, Energy and Communications and acts as the national contact agency for the EU Framework Programme for Research and Development. VINNOVA is also the Swedish government's expert agency within the field of innovation policy. |
| VQ | Vector quantisation |

# Acknowledgements

The writing of this doctoral thesis has been a huge undertaking, a Herculean task. The path which lead to the completion has been very interesting and enjoyable in every context. Perhaps o of the most interesting parts was to (for a while) ”leave” my own domain of computer science/informatics and computer engineering, in favour of acquiring deeper insights into the natural sciences, in this case biological sciences such as ecology, botany, forestry etc. In addition deeper insights have been acquired in behavioural science especially the neural academic field: information processing sciences like for example psychology. In the context of this thesis, this field deals with, among other things, human information processing for use in decision making. All this was then synthesised together with the computer science into this thesis.

Although it was hard and lonesome at times, I do believe that a thesis work of this magnitude is beneficial over time. It shows that concentration and mental focus can be kept during long periods, even for several years. I believe that this is what it takes to become an independent researcher who is able to manage projects of his/her own, and to independently apply for research grants, and to be able to run their own opinions and to provide rational arguments in any debate, and to be able to better act as a manager in a business or academic context etc etc.

First of all thanks to my family who have been supporting me day in and day out, namely my wonderful and beloved wife Merat and my intelligent and fantastic children Benjamin and Amanda. You are always in my thoughts! Thanks to my admirable and beloved parents Göte and Margot Nyberg, who have supported me in every context during my whole life. I love you all!!

Naturally I give thanks to my director-of-studies and supervisor Professor

Narendra K. Gupta (Edinburgh Napier University), my supervisors Prof. Mark S. Dougherty (Dalarna University) and Assoc. Prof. Siril Yella (Dalarna University) who all have proved to themselves to be excellent inspiring supervisors. Thanks to the independent panel chairs Dr. David Binnie and Prof. Brain Sloan, who chaired my progress reviews and encouraged me by interesting questions, twice a year.

Thanks to Jan-Erik Lundh project coordinator at the Swedish Transport Administration (who assisted me during some very hot summer days in the field experiments on the railway embankments), Prof. Anders Lindström, Assoc. Prof. Johan Håkansson, Marianne Vemhäll, Assoc. Prof. Eva Stattin, Dr. Daniel Wikström, Dr. Alam Moudud, Rickard Stridbeck, Ola Nääs, and Prof. Kenneth Carling. Thanks to Michael Nilsson, Alexander Karlsson and Nils Johansson (from the Graphic Technology subject) who have provided me the photo laboratory and photo equipment as well as their knowledge in photography. Thanks to the head of the department Anders Forsman, and the head of the academy Dr. L. Westman (until 2013) and Dr. G. Boklund (2013 - 2015). Thanks to Pär Douhan for all the fruitful discussions. Thanks to Calle Norman and Ulf Modin Research Engineer in Materials Science for supporting me with hardware solutions. Thanks to my colleagues in the departments of informatics and computer science.

Thanks to my dear friend Pär Nordahl who ultimately succeeded to motivate me for taking up systematic training again (after a down period of like 10-15 years). The training has certainly been supporting this work and I have kept this brilliant quote in my memory: "*Those who think they have no time for bodily exercise will sooner or later have to find time for illness*" (E. Stanley, Earl of Derby, 1826-93). Now I know this is true. Thanks to my friend Martin Olsén who (unknowingly) by the act of starting academic studies himself (computer engineering, 1996-1999) was the one who actually got me motivated and inspired me for taking up an academic career myself (which I started in 1998). Martin's quote: ”Hur svårt kan det vara!?” (Eng. ”How hard can it be!?”) has been constantly inspiring and will be a driving force in the future. Anything can be achieved if you are focused!

Thanks to all family members and friends who have encouraged and supported me during this work. Among those I would like to mention Pari, Zohreh, Ghazal, Hannah, Inga-Britt Nyberg, Stefan J. Eriksson and Anna Söderberg, Tommy Perlström, B-G Perlström, Staffan and Camilla Molin.

# Part I

# Introduction

# Chapter 1

# Introduction

Railway maintenance typically involves the following activities: inspections, extend lifetime of worn tracks by rail grinding, catenary maintenance including wire replacements, stabilising the tracks by packing (tamping) the track ballast, track including turnout replacements, ultra-sonic testing, repairing/replacing fasteners, joints, rails and sleepers, periodic measurements, signal repairs, vegetation management, measures against slippery rails, snow removal etc etc.

This thesis is dealing with subsets of the vegetation management part focusing on the *inspection and measuring of vegetation* serving as a maintenance planning instrument.

Subcontracting railway maintenance activities is not a trivial matter. The Swedish Transport Administration (STA) invites companies to make competitive bids for specific maintenance periods, involving various activities.

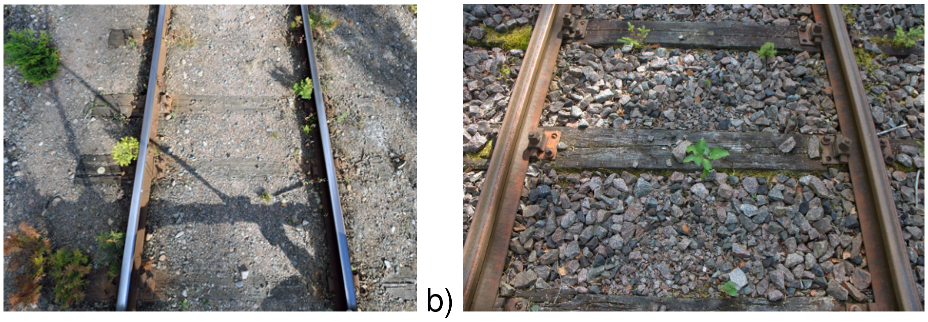
a)  b)

Figure 1.1: Growing vegetation on the trackbed a) Vetlanda, Sweden, and b) Falun, Sweden

With regard to ***vegetation management***, maintenance subcontractors find it extremely difficult to provide an estimate (often speculative) of the extent of rooted tree seedlings, herbs, and moss (see examples in figures 1.1a and 1.1b). The extent of vegetation is related to the workload that has to be put in by a railway maintenance subcontractor (henceforth known as a supplier of maintenance).Reliable information about the actual state of vegetation is often not available; thus, maintenance actions are carried out by subcontractors on a periodic basis irrespective of the condition. This wastes resources and results in the unnecessary use of herbicides. It is important to reduce the amount of herbicides used to fight vegetation along railways for environmental reasons.

Through our contact network, which consists of five railway maintenance subcontractors and the STA, we identified a common interest area of automated detection and characterisation of vegetation as a part of vegetation control (also called vegetation management). This offers scientific and commercial opportunities, because, currently, no system exists that deals with the problem in its entirety.

The track areas in consideration in this thesis (seen from the centre of the tracks) are the upper parts of the trackbed and the slopes of the embankment, not including the drainage ditches, see figure 1.2. Legend for figure 1.2; A represents the trackbed, B represents the embankment, C represents the slope of a cutting, D represents the drainage ditches, E represents the rails. The upper part above the dashed line in figure 1.2 represents a railway embankment, which carries rail traffic above the natural topography, i.e. over an area of low ground. The lower part under the dashed line in figure 1.2 represents a railway cutting, whether through soil or rock, carrying rail traffic below the natural topography, i.e. below an area of higher ground.

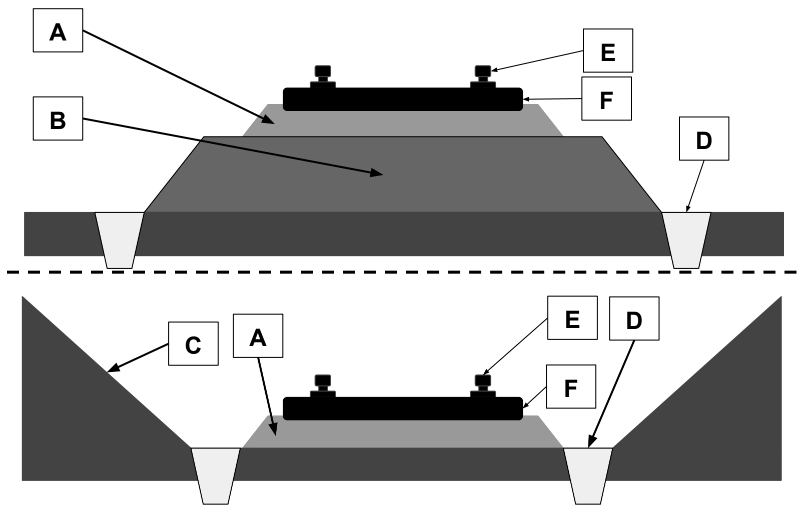



Figure 1.2: Overview of the track area

## 1.1 Problem Description

The vegetation problem is complex and includes several secondary problems. The basic problem is that there exists little, or no knowledge about the current vegetation state on properties that are part of the railway infrastructure. None of the involved actors, namely the subcontractors, track-owners and the Swedish Transport Administration, really know the current state out there. Often the state of the vegetation only becomes apparent when maintenance actions need to be carried out on-site; for example, inspections, mechanical harvesting, and the spraying of herbicides. By the time these actions are carried out, the procurement contracts have already been signed.

Vegetation growth along railways is often extensive (see examples on figures 1.3a and 1.3b); thus, maintaining an area free from such vegetation as weeds, shrubbery, and trees, is a constant struggle against nature (i.e the natural growth process). The STA's struggle with vegetation on and along railways is mainly related to ***safety*** (Banverket, 2000) (Trafikverket, 2012a) (Banverket, 2001) (Banverket, 2005b), both for passengers and their staff.

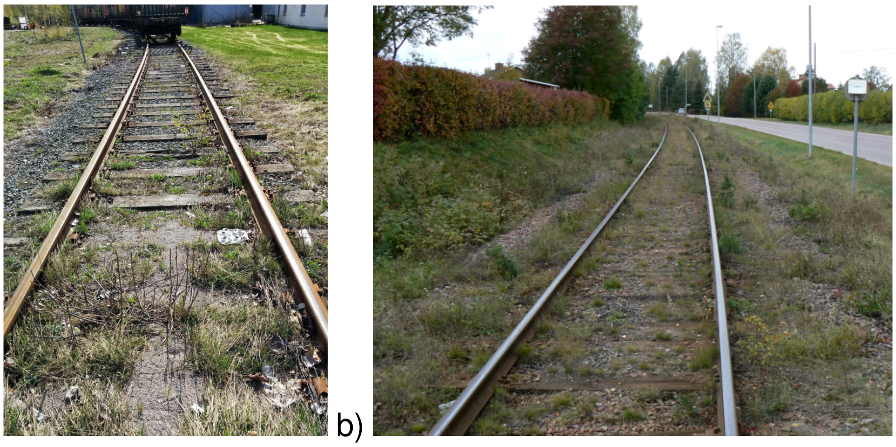
a) b) 

Figure 1.3: Vegetation growing on the trackbeds: a) Borlange, Sweden, and b) Mora, Sweden

If vegetation is permitted to grow uncontrolled, maintenance activities and operations will become a difficult, if not an impossible task. The major impact of uncontrolled vegetation is economic Scrivner (2004), because unmanaged vegetation accelerates the deterioration of every component in the railway infrastructure. Unmanaged vegetation and the use of herbicides along railways poses potential problems as follows:

1. The presence of vegetation threatens the safety of personnel inspecting the railway infrastructure, who may fall to the ground by slipping, or by tripping.

2. Vegetation that covers the railway tracks and trackbed makes it difficult for railway inspectors to detect potential errors along embankments, cuttings, etc.

3. A reduction in the elasticity of ballast could result from frost breaks caused by the presence of roots; this could contribute to potential land slides. The properties and composition of different types

of soil, groundwater conditions and topography (height differences) all determine the stability of a slope. The natural erosion process causes steep areas and slopes to change in order to find an equilibrium. Factors that alter the balance can trigger landslides. Such factors may include changes in soil water conditions, human impact on nature, erosion and uplift. Over time, therefore, stability conditions can change. Landslides occur through breakage along a sliding surface of the earth. The soil layers above the sliding surface are affected by driving forces, and by counteraction. Before a landslide, these forces are in equilibrium. However, a disturbance of the equilibrium can act as a trigger. The equilibrium can be disturbed by: 1) increased load, 2) reduced counterweight 3) deteriorated strength of the soil Caragounis (2014).

4. Growth of vegetation in the track indicates clogging of the ballast and a lack of adequate track drainage Chandra and others (2008, p.355). Hence, vegetation can lead to poor drainage of the embankment. Water is prevented from being diverted from the railway embankment; instead, it remains in the embankment body. It has been found that, during extremely wet winters, high hydrostatic water pore pressures can arise in grass-covered areas. This applies primarily to older railway embankments that contain clay (Briggs, 2010). (Scott et al., 2007) investigated seasonal variations of hydrostatic water pore pressures within embankment slopes. They found that they are controlled by climatic conditions, which are amplified by the effects of vegetation growth.

Figure 1.4: Collapse of railway embankment

5. Vegetation roots have a beneficial effect on strengthening the soil. However, in terms of slope stability, vegetation also have some negative effects.

    (a) (Scott et al., 2007) investigated seasonal deformation associated with the seasonal variation in pore pressures between the summer and winter months. During the summer, vegetation growth causes a pore pressure reduction, resulting in a downward movement of the slope surface. During the winter, the vegetation is inactive, which results in the rehydration of the soil, thereby causing swelling. In addition, the tendency for localised strains in clay fill results in the net downwards and outwards movement on preferred shear surfaces. This can lead to the development of a ”progressive failure mechanism” and eventual deep-seated failure.

    (b) During the course of a six-year period, (Smethurst et al., 2012) investigated seasonal cycles of soil water content, which led to shrinking and swelling in clay soils. This swelling and shrinking phenomenon often occurs in older railway embankments, and can contribute to strain-softening and the progressive failure of an embankment.

    (c) (Briggs, 2010) analysed a grass-covered cut slope made up of London clay (a stiff bluish type of clay), monitoring pore water pressures and water content over the course of one year.

He concluded that vegetation had caused a large cyclic change in effective stresses within the major drying zone (the top 1m depth of the profile) through a winter-summer-winter cycle. Recent evidence from both centrifuge model tests and numerical analyses of clay slopes suggests that cyclic stresses thought to be representative of those induced by vegetation can cause strain softening to occur, starting from the toe of the slope. Over a period of several years, these cyclic stresses can induce progressive failure.

6. Safety may be compromised because of visibility problems on bends and on level-crossings. For example, vegetation may obscure a vehicle's view of approaching trains from the vehicle's position at a level-crossing, as stated in one accident investigation report(Railway Accident Investigation Unit, 2010)

7. Vegetation becomes combustible during periods of hot and dry weather. In general there are two ignition sources: humans and lightning. Fires can damage the railway infrastructure.(Eddowes et al., 2003, p. 32)

8. Vegetation, including fallen leaves, on the rails makes it slippery and thus increases the braking distance of trains (Eddowes et al., 2003, p. 24). For this reason, so-called leaf buster trains (see figure 1.5) are used to blast the tracks with high-power jets of water. In addition a sandy paste (Sandite) can be applied to increase friction on the tracks.

Figure 1.5: Leaf buster train (Network Rail UK)

9. Concealed or occluded railway signals

10. Fallen trees and branches, often as a result of storms, or inadequate vegetation management, and blocked or damaged railway infrastructure can lead to short circuits in the overhead contact system. For example, (Russell et al., 2007) investigated momentary interruptions and sustained outages caused by vegetation intrusion.

11. Damage to the rolling stock; trains can be scratched or damaged by vegetation

12. Pathogens and vermin tend to propagate in areas of weeds, leading to complaints from nearby residents and farmers Hayakawa (2007).

13. Excessive use of herbicides in controlling vegetation poses serious threats to the environment. A reduction in the amount of herbicides used is desirable if we are to reduce their negative impact on the environment. Even though some types of herbicides are claimed to be less harmful than others, the existence of numerous environmental organisations and greater environmental awareness among the public make it hard to argument in favour for herbicides. For example, in a survey conducted by the International Union of Railways (UIC), Jan Skoog, Environmental Coordinator, Strategic Department at the Swedish National Rail Administration stated that, *”Concerns for groundwater protection and public opinion has led to the decision not to use chemicals where other possible alternatives are available” (Below et al., 2003, p. 104)*. Together with the desire to cut the costs of vegetation control, this has motivated several railway companies to start various activities to reduce the amount of herbicides used (Below et al., 2003, p. 5, 96 ).

With regard to the problems of using herbicides in vegetation management (see no. 13 in the list above), it is not only unwanted vegetation (weeds) and pathogens that thrive on railways; other animals and vegetation also enrich the fauna and biodiversity that are often present (Below, M., 27-28 April 2011, *Personal interview*). The danger of vegetation growth (e.g. on an embankment) does not necessarily lie in the few visible tufts of grass, or a couple of herbs. The real danger lies within the fact that vegetation grows slowly. This is most likely one of the primary reasons why vegetation management decisions get postponed again and again, year after year. This is probably because these problems never seem to be acute, one can always wait another year. By having this kind of attitude, it is possible for 5, 10, 20, or even 30 years to have ”suddenly” passed, at which point the problems have become acute. It might be to late for chemical, or mechanical vegetation methods of control. Indeed, the only solution left is to replace an embankment’s ballast, which is very expensive. (Also see section 3.2)

### 1.1.1 Procurement of Railway Maintenance

The procurement process of goods, services and contracts made by the STA (Trafikverket) is regulated by Swedish law. These are based on the Europeean Union (EU) Procurement Directives. In essence it means that all suppliers have to be treated in a similar and non-discriminatory way. STA must carry out all procurements in an open way, i.e., one that is public, so that every supplier can access the procurement process and anyone can view it.

The general procurement process can be seen in figure 1.6 and in the sequential list below (Trafikverket (Swedish Transport Administration), 2015) where each box in the figure is described.

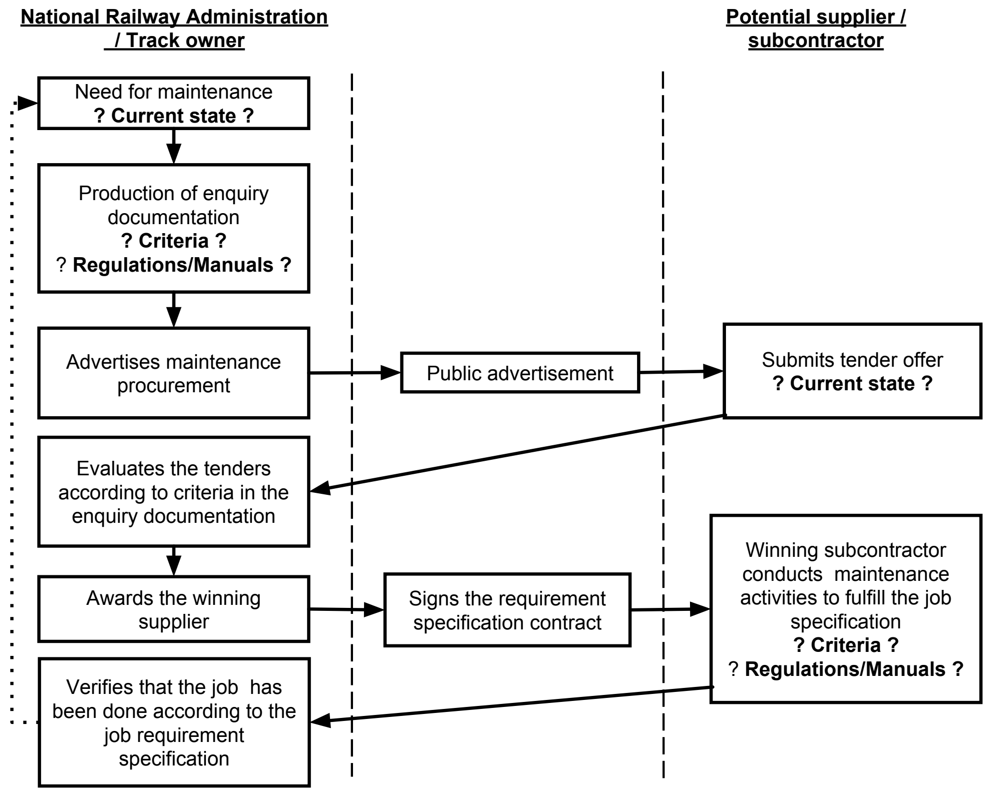



Figure 1.6: Maintenance procurement process

1. First, a need for maintenance arises within STA. This need will then be defined in *enquiry documentation*.

   - **Question**: Does the national railway administration (in the case of Sweden, the STA) have an exact knowledge of the current state, extent and (spatial) location of vegetation? If not, how did the need for maintenance arise?

2. Next, the development of procurement takes place. The enquiry documentation describes what is to be procured, what requirements are placed on the supplier and the subject of the procurement, e.g. vegetation management. It also incorporates how the tenders will be evaluated.

- **Question:** What documents (e.g. STA standards, regulations, manuals etc.) are used to develop and back up the list of criteria, to which the (winning) supplier must conform? Are the documents up to date? If the criteria makes references to and are based upon, for example, a regulation (e.g. (Trafikverket, 2012a)), or a manual (e.g. (Banverket, 2000)) to which the winning supplier must conform, then how objective and precise are the instructions in that regulation or manual?

3. A request for suppliers to tender is advertised.

4. The interested suppliers check out the advertisement and make an estimate of the price for carrying out the required job. The supplier sends in their tender.

    - **Question:** Do the suppliers have knowledge of the current state, extent and (spatial) location of vegetation? If not, how can they make an estimate of the workload that has to provided by their company?

5. The STA evaluates the tenders in accordance with the evaluation criteria set out in the enquiry documentation.

6. The STA notifies all those who made tender offers. They then set out which supplier(s) has (have) been awarded the contract.

7. The contract is signed by the STA and the winning supplier(s).

8. The winning supplier(s) carry out their work according to the contract. This action is iterated within the action listed in no. 9 during the contract period.

    - **Question:** If the contract contains job requirements that refer to STA regulations or manuals (see the development of these in list no. 2), then is it possible to objectively follow the instructions in that regulation or manual?

9. The actual work (see list no. 8) and what is stated in the contract is continuously followed up by the STA during the term of the contract.

10. When the contract has expired, this process starts over again from list no. 1.

In a report from (Eddowes et al., 2003, p. 24) to the UK Rail Safety and Standards Board (RSSB), climate change was investigated. Climate change, which implies increased temperatures, heavy rainfalls, increased wind speeds and so on, together with increased vegetation growth in the railway infrastructure, may lead to increased production of (plant) biomass because of the longer growing season (Eddowes et al., 2003, p. 24).

A rise in temperatures, increases in the frequency and length of dry weather, and an increase in lightning strikes, also increases the risk of fire developing in vegetation (Eddowes et al., 2003, p. 32). On the positive sides for vegetation, the stabilising properties of plants growing on embankments is highlighted.

The **first step** in vegetation management is to **detect** and **characterise** the **vegetation**. This is the base for maintenance decision making. This step is currently solved by manual inspections along the railway made on foot, or by civilians who contact the STA and make them aware of certain conditions. In some cases, such as periodic maintenance and herbicide spraying, no inspection are carried out (Lundh, J-E., 28 April 2011, *Personal interview*).

## 1.2 Manual Inspection Routines

Currently, daily maintenance activities aimed at dealing with vegetation along railway tracks are carried out manually. This section describes how a general (safety) inspection is carried out, and outlines the regulations to which a inspector must conform. Typically, human inspectors are employed by subcontractors, who are responsible for the maintenance of certain railway areas. These inspectors walk along the railway tracks and judge for themselves the extent of the vegetation and its condition.

In cases of poor condition, the inspector recommends further maintenance actions. Note that such decisions are made intuitively by the inspector and are largely dependent on the Swedish Transport Administration regulations. Requirements regarding vegetation are regulated in BVH

807-series(Trafikverket, 2012a) (Banverket, 2005b) (Banverket, 2005a) (Banverket, 2005d) (Banverket, 2005e) (Banverket, 2005c) , and 827-series(Banverket, 2000)(Banverket, 2001) among others, such as (Trafikverket, 2012c).

### 1.2.1 Administrative Regulations

The Swedish Transport Administration (STA), is an authority that sits under the Swedish government, and has the overall responsibility - sector responsibility - for the Swedish railway (and road) infrastructure, i.e., the rail transport system. In this context, the STA performs both administrative and productive operations. Requirements regarding vegetation are regulated in Swedish Rail Administration documents 807- and 827-series, including important regulations and handbooks. These handbooks are to be used by any maintenance subcontractor hired by the STA.

The following part of this section describes the framework of some of the regulations to which any subcontractor (wanting to perform maintenance on the railway infrastructure) has to commit themselves. As this work is about monitoring vegetation, the chosen administrative regulations address this area.

#### Types of Inspections

In Sweden and to a large extent elsewhere around the world two types of inspections are deployed:

1. *Safety Inspections*
2. *Maintenance Inspections*.

Inspecting vegetation is a subset of both types.

The purpose of *safety inspections* is to verify that there are no anomalies, as well as to identify and estimate the deterioration of inspected objects. As a result, STA, or its subcontractors, are able to prevent or eliminate defects that can lead to accidents or incidents relating to trains, electrical safety incidents, staff working injuries, accidents involving third parties, operation disturbances, and environmental incidents.

The purpose of *maintenance inspections* is to serve as a foundation for planning actions in the medium to long term, equalling six months to three years Banverket (2005b, p.5). This is to meet the requirements of function, optimal tech-nical and economic lifespan.

### 1.2.2 Estimate of Inspection Costs

A general safety inspection (see section 1.2.1) carried out on Sweden's 14 000 km tracks by walking along the trackbeds at an average speed of 3.5 kilometres per day, would have resulted in $14000/3.5 \approx 4000$ days, equalling approximately $4000 * 8 = 32000$ man hours (which in effect is about 11 years of full-time work for one man). In 2011, the cost of a railway inspector (including necessary safety precautions) was around 100 GBP per hour and about 800 GBP per day. The resulting cost for inspecting all 14000 km railway tracks results in: $32000h * 100\,GBP/h = 3\,200\,000\,GBP$ $\approx 4\,900\,000\,USD \approx 3\,600\,000\,EUR \approx 36\,000\,000\,SEK$. Note that this general safety inspection covers far more then just inspecting the vegetation, see section 1.2.1. Based on the numbers above, the average cost per kilometre would then equal $800/3.5 \approx 230$ GBP/km. (The average speed and costs are based on Sjöblom, T., 1 March 2011. *Personal interview)*

A field trial led by Jan-Erik Lundh at the STA in 2011 aimed to investigate the methods of manually inspecting/monitoring vegetation in the drainage ditches beside the trackbed. It took about 40 minutes for an inspector to inspect/monitor one kilometre. Hence, in this case, the average speed resulted in progress of 1.5 km/h, or 12 km/day. At this speed, and at an hourly cost of 100 GBP, the average cost per kilometre would then be $800/12 \approx 70$ GBP/km. (Based on Lundh, J-E., 28 April 2011. *Personal interview*)

### 1.2.3 Disadvantages with Manual Inspections

Relying on human inspectors to walk along railway tracks is time consuming and expensive. In addition, manual inspections require skilled and trained staff. Maintaining a consistent quality standard is also difficult, because different inspectors may interpret the regulations in different ways, or they may have different opinions on how inspections should be carried out. Consequently, inspection results are highly subjective, depending on the experience and knowledge of the person carrying out the inspection. This subjectivity can result in uneven quality.

The notes made by human inspectors should be dependent on the STA requirements stated in the regulations and handbooks. However, variations in inspection notes do not just result from the different opinions of inspectors. Subjectivity is also the result of inconsistent administrative regulations, which have very few measurable criteria.

Two examples of administrative regulation texts that show degrees of subjectivity:

> *"Verify that no interfering vegetation is present in the track area"* Banverket (2005b)

How should the words *"interfering",* and *"present",* respectively, be interpreted by an inspector?

> *"At level crossings: Verify an adequate field of view, and that road signs are not obscured*"(Trafikverket, 2012a)

Likewise, how should the wording *"adequate",* and, *"field of view",* and *"obscured",* respectively, be interpreted by an inspector?

Apart from the problem of intuitive manual inspections, reliable information about the current state of vegetation along railway tracks is not readily available; thus, maintenance actions are carried out manually on a periodic basis irrespective of the condition, which is very slow and time consuming. The fact that some high-traffic lines have to be inspected more frequently, as well as other issues such as maintaining an even quality standard, makes high costs a serious problem.

In efforts to chemically control vegetation, woody plants, herbs and grasses along railways, Glyphosate (commercial name: Roundup Bio) is often used(Torstensson, 2001). Glyphosate does not affect conifer trees that much. Whilst it will kill new shoots, the plant itself will continue to grow.

Equisetum arvense (horsetail) is also not affected by Glyphosate. The substantial use of herbicides to fight vegetation along railways poses serious environmental problems.

Detailed railway maintenance information, e.g. previous inspection protocols, often resides with the subcontractor who carried out the inspection, and thus the information becomes inaccessible over time.

The STA currently performs periodic measurements by video recording the track and surroundings using a measurement wagon (called a STRIX). These video clips may be viewed by subcontractors; however, the personnel watching the video clips still have to make the decisions regarding the maintenance measures to be taken. Information contained in the video clips is not represented or stored in any other form (except for the GPS-based positioning information on the STRIX wagon carrying the video camera). Thus this kind of video information is static.

To deal with the consequences of climate change, AEA Technology PLC in the UK has put forward a proposal to the RSSB which makes the following research recommendations concerning vegetation: 1) Develop improved general awareness of anticipated impacts and timescale, 2) Develop a programme to monitor changes, 3) Implement revised vegetation management practices to meet new growth characteristics (Eddowes et al., 2003, p. 54). The second research proposal legitimises work carried out in this thesis.

In particular, the described disadvantages of manual inspections has led to the **suggestion of an automated approach** in order to support or replace manual inspections (see Section 1.3).

## 1.3 Proposed Automated Solution

Initial enquiries, including a literature review and interviews, show that there are no automated systems for monitoring vegetation along railways in Sweden. There is a strong need for such systems developed for this purpose[1] [2]. Automating the process of *detecting* and *characterising* vegetation along railways could be accomplished by replicating human vision, using machine vision techniques. More than 95% of the information that humans perceive is optical in character (Flusser et al., 2009), so human expert knowledge could be extracted and used in the contribution of developing the foundation of such a system. A machine vision solution could also be used for *identifying* common species of woody plants (which are more problematic than herbs) as well as endangered species that grow on the railway embankment.

Proposed data sources include on-board sensors (e.g. cameras, and GPS). Use could also be made of remotely sensed data. The information output of the automated inspection process will serve as a knowledge base for maintenance decisions, i.e., planning when and where to carry out vegetation control/management. The sensors should optimally be available in the commercial market at a reasonable price so that most subcontractors can afford to use them.

The system should be robust, and fault tolerant. For example, it should be possible to perform monitoring even though a camera is slightly misplaced, or set up incorrectly. Thus, the system should be able to *sense the environment by itself using sensors*, i.e., the system should be able to locate objects like railway embankments, rails, sleepers, rail fastenings, base-plates, and so on. This would also enable the same system to be used for monitoring activities other than vegetation monitoring.

## 1.4 Research Questions

The research questions (RQ) that address the problem of monitoring the state of the vegetation along railways are listed as follows:

[1] Enquiries made through interviews with Trafikverket and local railway maintenance subcontractors, as well as written application response from domain experts at VINNOVA .

[2] Acknowledged by local railway maintenance subcontractors working in the railway maintenance area, Trafikverket (STA) as well as domain experts at VINNOVA, 2009.

RQ 1: How are railway inspections carried out with regard to the assessment of the extent of vegetation, and what methods are used for measuring vegetation along railways?

RQ 2: How is the extent of vegetation measured?

RQ 3: How reliable are human visual estimates when assessing the extent of vegetation?

RQ 4: How can the extent of vegetation be measured by making use of machine vision, machine learning and statistical inference?

RQ 5: How can woody plants growing on railways be recognised using machine vision and machine learning? Woody plants are problematic because legal herbicides are not completely effective in killing them (see section 1.3); thus, it might be feasible to mechanically harvest them if the maintenance administrators knew their spatial position along the railway system.

RQ 6: How do measurements using machine vision and machine learning correlate with human visual estimates?

## 1.5 Aims and Objectives

Based upon the the proposed solution (see section 1.3) and the research questions (see section 1.4) the following aims and objectives are defined.

### 1.5.1 Aims

The aim is to investigate and develop a potential automated system for the detection and characterisation of vegetation on railways in order to:

1. Reveal the state of vegetation on railways through a geographical overview of the railway infrastructure for subcontractors, track-owners and administrators (e.g. national railway administration), thus synchronising the common knowledge-base of a particular section of embankment for, for example, outsourcing maintenance activities.

2. To recognise woody plants that are most likely to grow on railways. By knowing where these are in the railway system, it might be feasible to mechanically harvest them.

3. Support maintenance decisions for subcontractors, track-owners, and administrators.

4. Enable a shift from periodic maintenance to condition-based maintenance, thus reducing costs and saving resources. In addition, this may prevent the overuse of herbicides.

### 1.5.2 Objectives

The main objectives pursued in order to answer the research questions are:

1. To investigate problems that arise when carrying out manual vegetation inspections along railways as well as to investigate current-day vegetation control/management activities.

2. By using sensor inputs, to propose computational methods and algorithms that detect, measure and characterise vegetation and to map it temporally and spatially.

### 1.5.3 Effects

The effects of automating inspections with respect to vegetation are as follows:

1. Improved track safety and reliability, as opposed to current individual subjectivity.

2. The initialisation of a programme to *monitor changes of state* along the tracks, leading to:

    (a) improved procurement of vegetation management (see flow chart in section 1.1.1). Both the railway administrators and potential subcontractors know the extent of vegetation beforehand. After the contracted work has been carried out, it will be easier to follow up by comparing the previous recorded state.
    (b) increased knowledge of where endangered species are to be found along the tracks.

3. A dynamic condition-based maintenance (CBM) approach is achievable, as opposed to the day-to-day static periodic maintenance that currently takes place (see section 2 for further information about maintenance).

4. Avoid unnecessary delays through improved decision making.

5. Implementing a CBM can reduce the amount of herbicides used, thus making less negative environmental impacts and lowering cutting costs for vegetation management using herbicides.

## 1.6 Research Methodology and Design

Based upon the presented aims and objectives, this section describes the research methodology used in this thesis. Also presented is a summary of how data was collected and how it was analysed.

### 1.6.1 General Methodological Approach

Overall, this thesis emphasises both the scientific method and the engineering design process (see appendix I). Several interdisciplinary studies were incorporated. At its base are the academic fields of computer engineering and computer science, or informatics, and focuses on pattern recognition and machine vision. It also heavily involves ecology, forestry and botany, where the methodology of describing using measurements is in focus. Plants are the main subject to be measured; thus, biometry is in focus. Biometry is the science of measuring and statistically analysing biological data. The fields of medicine and psychology are emphasised, in conjunction with behavioural science. Foremost, they are referred to in terms of measuring human reliability (levels of agreement among raters/ assessors/judges/observers). This could involve, for example, several raters visually estimating/assessing plant frequency.

The overall research process was conducted in an iterative way and involved literature reviews. The data were collected through measurements, questionnaires, and interviews.

The underlying research methodology used emphasises quantitative research methods, foremost by observation in experimental studies and surveys (see figure 1.7). Reference has also been made to experiments and surveys as they form the basis of this thesis.

Aspects of qualitative methodology were used whenever subjective interpretations were needed for a deeper understanding, e.g. interviews with experts from a particular domain.

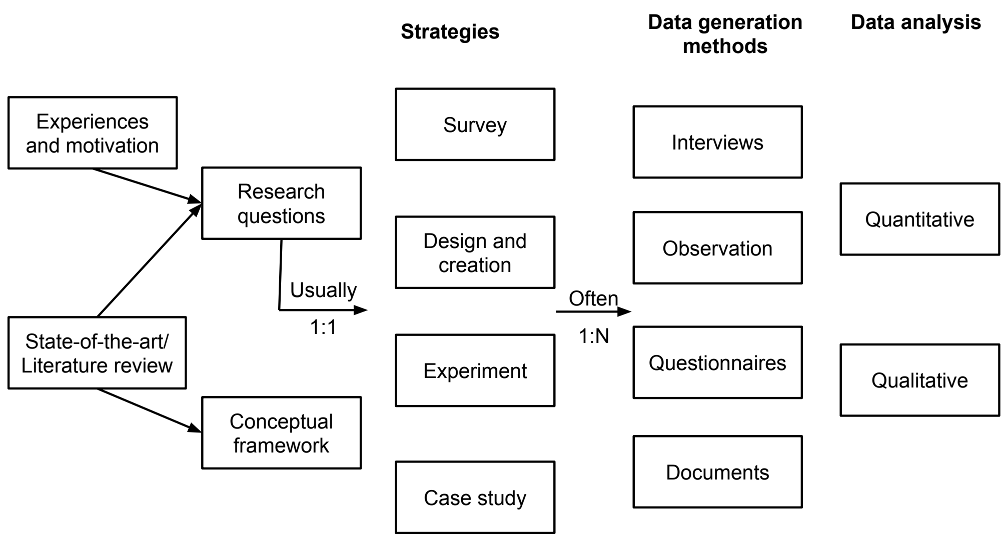



Figure 1.7: General Research Process (Modified and redrawn from Oates (2006))

The investigations in this thesis included both *controlled experiments* (for example, in a photo laboratory) and *field experiments* in an uncontrolled outdoor environment along railway trackbeds. In general, the following components formed part of any experiment (Cox and Reid, 2000, pp. 19)(Hurlbert, 1984):

1. Experimental hypothesis

   One or several well-formed questions or research hypotheses, e.g. speculations about some underlying process, or phenomenon that confirm or explain earlier findings.

2. Experimental design and execution

   Description of the nature of the experimental units, the number and kind of treatments to be applied, and what will be measured in response to the experimental units. The experimental design specifies how treatments are assigned to the experimental units, the number of experimental units (replicates) receiving each treatment, the spatial arrangement of experimental units, and the temporal sequence in which treatments are applied, and measurements made on the different experimental units.

   A goal during the execution of the experiment is to: 1) avoid the introduction of systematic errors (bias), and 2) to minimize random errors. It is very important to get it right in this phase, because any errors made at this point often mean that the experiment needs to be repeated. This can be very costly (in terms of time and money). Sometimes, an experiment cannot be repeated simply because the objects to be studied are no longer there.

3. Statistical analysis

   The data are analysed by making use of objective statistical methods. This does not necessarily imply that the result is objective, but that the statistical methods are. Hence, if someone else using the same data was to repeat the statistical methods, the result would be the same.

4. Interpretation of the analysis results and conclusion

The experimental design in this work are either randomised or systematic. This refers to the sampling design. A randomised design means that the observed units have been chosen by a simple random process out of the population. In the case of systematic design, the starting point (e.g. on the railway track) has been randomised and the subsequent points has been predetermined by intervals in time or space (e.g. sampling occurs every 5th second, or every 10th meter). This method is also called an every $k^{\text{th}}$ systematic sample (Cochran, 2007, p.205 ). The statistical risk of errors in the systematic design occurs if the environment is periodic. The chosen and adapted design for each experiment is explicitly described at the point where the experiment is presented.

Artefacts such as models, algorithms and prototype implementations were created during this work using crucial elements of the experimental strategy. Foremost, alongside the design and creation strategy, a problem-solving approach was taken. The design and creation strategy has it roots in design science (see Oates (2006, pp. 108-124), March and Smith (1995)). The scientific study of artificial entities was introduced by Herbert Simon in his book "The sciences of the artificial" (Simon, 1996), first edition published in 1969. Here, Simon put forward arguments in favour of a science of design of artificial entities in parallel and/or supported by natural science. The aim of natural science is to understand and explain natural phenomena, whereas the aim of design science is to develop ways to achieve human objectives (March and Smith, 1995). In this thesis, the overall objective is to investigate and propose methods for quantifying vegetation along railways using camera-based sensors.

Whenever *interviews* were used as a method for data collection, the main purpose was to develop a *deeper understanding of the domain* in question. Therefore, interviews in this work have been conducted as *unstructured interviews* (Weiss, 1994, pp. 207)(Fontana and Frey, 1994, pp. 365–368), or as *short case study interviews* (Yin, 2014, pp. 109).

Unstructured interviews are characterised by the interviewer having a loose plan in mind concerning the goal of the interview and this goal is used as a guide through the discussion. Neither the interviewer nor the respondent are restricted by, for example, predefined questions. Questions in the unstructured type of interview are open-ended, i.e., the questions relate to what, why, how, describe, explain, and compare; they do not lead to a simple yes/no response.

The interviews were used as an important preliminary step before the development of more sophisticated investigations or surveys, and to support or enhance findings in the existing literature.

### 1.6.2 Research Design

This section describes how the chosen methods were applied to answer the research questions. In order to be able to answer the research questions in section 1.4, the work was divided into several iterative phases (see list below). The addressed research questions which were to be answered by some actions (in the list below), are denoted as <RQ*n* >, where *n* is the number of the research question.

**Phase 1 - Background Research and Literature Review** This phase involved several review parts:

1. A review of previous work and technical documents concerning *railway vegetation control*. These documents investigate how vegetation on and along railway embankments are currently detected and treated, as well as how manual inspection routines are carried out. This review addresses RQ1 .
2. A review of *machine learning and machine vision methods* are candidates for this work. This action addresses RQ4.
3. A review of methods for monitoring and measuring terrestrial vegetation. This review addresses RQ2

**Phase 2 - Data Collection**

1. Primary data were collected outdoors (denoted as *field experiments*) by acquiring digital images using camera sensors on different railway sites during the non-snowy seasons. In addition, primary image data were collected in a laboratory environment indoors. The purpose of the latter was to test the concept of recognising/classifying plants in a controlled noise-free environment, before doing the same in a noisy outdoor environment, i.e., on

the railway embankments/trackbeds. The collected data constituted the base for the upcoming analysis of investigations, wwhich attempted to answer RQ3, RQ4, RQ5, and RQ6.

2. Primary data were also acquired by putting survey questions to domain experts and laypersons about acquired images (above), or by observation on-site during the field experiments. The data were used to answer the question of reliability regarding human estimates of plant cover. This data also, as above, addressed RQ3 and RQ6

**Phase 3 - Analysis of Investigations**

1. *Manual inspections*: Investigations of laypersons or domain experts who assess the extent of vegetation were statistically analysed. Data acquired in phase 2, above, were used. This targets RQ3, by making comparisons between manual inspections made by different people.
2. *Machine vision inspections:* Here, the acquired data (from phase 2) were analysed and algorithms were developed. The algorithms were then implemented in the context of image processing, which underlies machine vision, and machine learning. This iterative development of prototype software also included testing. This step addresses RQ4 and RQ5.
3. *Comparisons* between human visual estimates and the machine vision/machine learning approach were analysed. This addresses RQ3 and again RQ6 but this time by comparing measurements made by human vs. machine
.

**Phase 4 - Discussion and Conclusion** Finally the results from the previous phases were discussed and suggestions for future work were made.

## 1.7 Scope and Limitations

### 1.7.1 Scope

The general purpose of this thesis is to investigate and propose methods for quantifying vegetation along railways using camera-based sensors.

This thesis covers the following topics: a problem description of letting vegetation grow along railway lines, a description of general maintenance methods (especially in the railway domain), an overview of those fields of botany and ecology that target the basics of plants and a description of how vegetation is quantified, investigations and descriptions of adequate machine-learning methods in conjunction with machine vision, investigations of human raters (e.g. maintenance inspectors), and agreements and the reliability (or trustworthiness) in their assessments.

As already stated in the opening paragraphs of chapter 1, the track areas in consideration in this thesis (seen from the centre of the tracks) are the upper parts of the trackbed and the slopes of the embankment, not including the drainage ditches, see figure 1.2.

Image data were collected at various field experimental sites in the county of Dalarna (Dalarnas län) during the summer months between 2010 and 2012, and in the county of Småland (Smålands län) during the summer of 2013 (see section 5 for detailed locations in Sweden). The locations were selected bearing in mind aspects relevant to safety and site availability. The specific experimental sites at the locations were always selected at random.

Survey data were collected from both domain experts and laypersons, where domain experts were considered as persons working in the railway domain and/or having any knowledge of how to assess the extent of vegetation, and for the layperson the opposite applies.

Typically, the respondents (i.e., the raters), were asked to assess the extent of vegetation either in the field (i.e., on the railway embankment), or by looking at images of a railway embankment. The survey used a bounded continuous type of response scale, where the respondents were presented with a continuous scale. For example, when assessing vegetation cover (i.e., estimating the percentage of plants covering a predefined area), respondents were asked to provide an answer between 0 and 100%. With regard to vegetation density (i.e., counting plants/clusters in a predefined area), respondents were asked to provide an answer as a natural number, or zero. $\mathbb{N}^0 = \{0, 1, 2, 3, ...\}$. The purpose of conducting the surveys is to assess the ability of human raters to assess the extent of vegetation and analyse whether if such assessments are reliable.

### 1.7.2 Limitations

The results from the outdoor field investigations should be handled with care, particularly when it comes to generalising them to other field layer classes (see section 5.2) and to other climate zones, where more investigations are needed.

The investigations in which laypersons and domain experts were asked to assess the extent of vegetation were not intended to to be (statistically) exhaustive; rather, they were to serve as a proof-of-concept indicator in conjunction with the cited literature. Few domain experts were available. If more experts had been available, fewer investigations could have been carried out instead.

## 1.8 Outline of the Thesis

In order to help the reader, a short outline of the thesis is presented in this section. The thesis consists of the following main parts:

1. **Background, applied strategy and knowledge base:** Consists of a description of the problem posed by vegetation growing along railway tracks and how it is dealt with by railway administrators or (often) subcontracted maintenance suppliers. Most essential is one of the first steps in maintenance: How is the state of vegetation along railways inventoried or monitored? Next, the strategy for solving the problem is presented. The research questions are based on the problem and form the foundation of what is investigated (for more information see part I). Furthermore, a review of maintenance fundamentals

is given, together with a discussion of how terrestrial vegetation is being measured by experts (see part II for further details).

2. **Data collection methods:** Part III contains the basics of why, when and how the data was collected. Instead of describing the data collection several times, data sets collected from the same temporal and/or spatial context were used later, at several places in the thesis. Thus, data collection methods are shown in one place.

3. **Humans quantifying vegetation:** Part IV describes the investigation of how effective humans are at estimating the extent of vegetation. Methods described in chapter 4 are used.

4. **Quantifying vegetation using machine vision and machine learning:** In part V alternative ways of measuring vegetation are presented. Foremost is the detection and quantification of vegetation by making use of machine vision. The detection and classification of common woody plant species using machine vision and machine learning are also presented.

5. **Conclusions and discussion:** The conclusions and discussion are followed up by suggestions for future work in part VI.

# Part II

# Technology and Literature Review

The ideas leading to this thesis have their origin in computer science/ engineering. However, they are also very much grounded within several other disciplines, such as biology in general, and botany or plant/vegetation science, ecology, and statistics in particular. Thus, this part of the thesis covers some of these areas. It should also be noted that, in several cases, literature is cited elsewhere. The motivation for this explicit choice is to use citations in a relevant context of reasoning. Thus, it would not have worked as well if they were only found within a technology and literature review such as this. This approach also enhances readability, because the reader does not have to go back and forth between the chapters.

In some contexts, the sources cited in this work may appear as old at first sight, especially concerning the domains of ecology, biology, botany or plant science and statistics, where citations from the 1930s to the 1960s are not unusual. However, these sources still hold true today and are very much in use in current day research.

A table of the Internet-based sources used during the technology and literature review is shown in Appendix K.

# Chapter 2

# Maintenance Strategies

The bottom line goal of this study is to support and improve the maintenance activities related to vegetation management (see also section 1.3). This chapter describes maintenance, together with any terms and definitions used.

Two general types of strategies concerning maintenance can be distinguished Yam et al. (2001) :

- *Corrective maintenance (CM), Run-To-Failure* (RTF), *break-down maintenance*, or *unplanned maintenance* all take place only when failure occurs, i.e., one waits until the equipment or material in question fails or breaks down; only then is it repaired. This may be a good strategy if the equipment or material is cheap, and the cost of repairing or replacing it is lower than keeping up a maintenance programme.
- *Periodic maintenance, preventive maintenance, time-based maintenance*, *scheduled maintenance*, or *planned maintenance* all take place at periodic intervals, regardless of the state or condition of the object in mind.

The European Standard EN 13306 BSI (2010) provides an overall view of maintenance (see figure 2.1). *Preventive maintenance* is broken down into *condition-based maintenance (CBM)* and predetermined maintenance. A *preventive maintenance* is carried out at predetermined intervals or according to prescribed criteria. The purpose is to reduce the probability of failure or the degradation of a functioning item.

A *predetermined maintenance* is a type of preventive maintenance and is carried out in accordance with established intervals of time or number of ”units” of use, but *without* previous condition investigation (Compare with CBM). A *scheduled maintenance* is a type of *predetermined maintenance* and is carried out according with a time schedule or established number of ”units” of use.

A *CBM* is, like the predetermined maintenance, a type of preventive maintenance and includes a combination of condition monitoring, inspection, or testing followed by analysis and subsequent maintenance actions. A *predictive maintenance* is a CBM carried out following a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item.

A *corrective maintenance* is carried out *after fault recognition* and intended to re-establish functionality.

A *deferred maintenance* is a delayed type of corrective maintenance and is thus not carried out immediately. The opposite is the *immediate maintenance* that is carried out without delay after a fault has been detected.

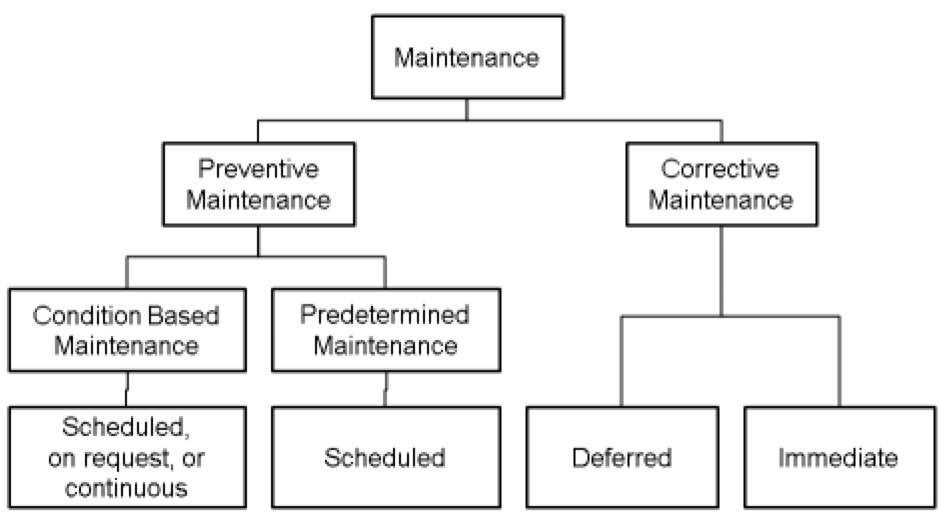



Figure 2.1: Maintenance - Overall view (Redrawn from BSI (2010))

With the rapid development of modern technology, products have become more and more complex whilst, at the same time, better quality and higher reliability are required. This brings the cost of preventive predetermined maintenance higher and higher, Jardine et al. (2006).

An often more efficient maintenance approach is condition-based maintenance (CBM). CBM is a maintenance program that recommends maintenance actions based on the information collected through condition monitoring. In general, maintenance actions are carried out each time the value of a given system parameter exceeds a predetermined value. CBM attempts to avoid unnecessary maintenance tasks by taking maintenance actions only when there is evidence of abnormal behaviours of a physical asset Jardine et al. (2006). Hence, if CBM can be effectively implemented, it will reduce the number of scheduled preventive maintenance operations and thereby reduce overall maintenance costs.

CBM approaches can be divided into two main categories: diagnostic and prognostic Wang et al. (2002). The diagnostic approach involves such steps as detection, isolation and the identification of faults when they occur. The prognostic approach attempts to predict faults or failures before they occur. Within both categories, three further types of approaches can be distinguished: statistical approaches, artificial intelligence (AI) approaches and model-based approaches Jardine et al. (2006).

A CBM program basically consists of three sequential stepsLee et al., 2004, see figure 2.2. The first step is to acquire data from the monitored environment. The second step is to process and analyse the gathered information. The third step is to make decisions based on the analysis.

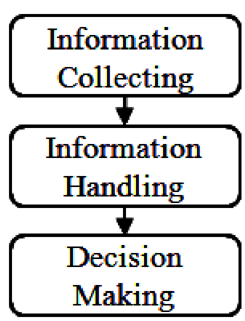



Figure 2.2: Steps in condition-based maintenance

The advantages of CBM have been reported by Khazraei and Deuse (2011); maximising equipment availability and machinery life expectancy, and found that it reduces overtime costs and decreases rework. CBM provides an analysis of failure causes and a more economical use of maintenance resources. However, the same authors reported that these techniques require specialised equipment and staff training, which could become expensive. With regard to vegetation management, the STA mainly conforms to scheduled predetermined maintenance, and more specifically time-based preventive maintenance, where for example the chemical treatment of railway embankments occurs on average every third year (Lundh, J-E., 28 April 2011, *Personal interview*).

For a more comprehensive review of maintenance terms, strategies, standards, and maintenance in general reference should be made to Khazraei and Deuse (2011).

## 2.1 Vegetation Monitoring

In the context of monitoring, the term *survey* often appears. A survey can be defined as a set of quantitative or qualitative observations using standardised procedures and within a restricted period of time, without any preconception of what the findings might be Goldsmith (2012); Hill (2005, pp. 552). A series of replicated surveys, denoted as surveillance, is here defined as monitoring. *Monitoring* can be defined as recurrent (regular or irregular) surveillance to determine the degree of compliance with a predetermined standard or degree of deviation from an expected norm. Monitoring is performed when one wants to test the effects of macro policy on a large spatial scale. Therefore, the design of data collection and the analysis of the same must be given careful consideration. Hill (2005, pp. 3-5). In the context of nature conservation, (Alexander, 2008) stated that, "Monitoring should be an essential and integral component of management planning: there can be no planning without monitoring and no monitoring without planning". This is very much a universal statement that complies to any maintenance and management process in any context, and applies particularly well for the purpose of this work.

## 2.2 Sensor Technology

In general, a sensor is a device that detects or monitors its surrounding environment and then records, indicates or responds to it in some way. One of the purposes of this thesis is to be able to characterise, detect and identify vegetation growing on railway trackbeds, and hence only sensors applicable for this purpose are mentioned. One of the most common approaches in using sensors to characterise vegetation is to make use of the electromagnetic spectrum. More specifically by capturing multispectral (or hyperspectral[1]) images at specific bands of frequencies across the electromagnetic spectrum. In conjunction with multispectral images, the the amount of light of a specified wavelength in the electromagnetic spectrum could also be measured by optical spectrometers. Multispectral and hyperspectral imagery has its applications in the fields of ecology, forestry, agriculture, minerals, oil and gas industry, oceanography and atmospheric studies

Depending on the purpose of the investigation it is often common to make use the ultraviolet light (approx. 10 to 380 nm), visible light (from violet around 380 nm up to red at around 700 nm), or infrared light (approx. 700 nm to 1 mm).

As stated in the proposed solution (see section 1.3), the chosen sensor(s) should optimally be available in the commercial market at a reasonable price so that most subcontractors can afford to use them. Hence the selection of sensors are therefore restricted by their price as well as availability. The choice of image sensors is also justified by letting those who collect and analyse data to actually look at it themselves and make their own judgements Dougherty (2013). In addition, regulations and manuals of national railway administrations are based on the outcome of visual inspections which can be mapped into the image and video clip domain. Another benefit of using image sensors (e.g. digital single-lens reflex cameras, DSLRs) is that they produce files according to common standards, such as RAW, jpeg etc.

[1]Hyperspectral imagery has much narrower bands (10-20 nm) than multispectral imagery

### 2.2.1 Sensor Technology for Detecting Vegetation

A nitrogen sensor (aka ***N-sensor***), like the commercial Yara sensor and others, has been the subject of some studies, including one by Fogelberg (2011). An N-sensor is typically used on crop fields by farmers who want to optimise the distribution of fertilisers on their fields. The sensor determines a plant's nitrogen demand by measuring the plant's light reflectance. By analysing this input, a certain amount of fertiliser will be distributed on the monitored crop field. Although it seems like many things are included in this one commercial system, it does have drawbacks. These include the fact that it is a black box, which means that no changes can be made to the system. In addition, no checks can be made on the way it computes the need for nitrogen, and other sensors cannot be coupled with the system without the approval of the company.

An example of a final output after evaluating a railway section with a Yara N-sensor can be seen in figure 2.3, where a darker green colour indicates a higher prevalence of vegetation. The resulting output, on a high aggregated level, serves a purpose for, for example, farmers who want an overview of their fields. However, it might not be a final off-the-shelf product for monitoring vegetation on railways simply because of the high aggregated result levels which cannot become any more precise. Furthermore, it is highly expensive, so it would take several years for such a system to give a return on investment for a railway maintenance supplier who specialises in vegetation management. An N-sensor measures electromagnetic light, so the same task can be accomplished by and controlled by other optical sensors, e.g. VIS and near infra-red (NIR) cameras, or, as described next, by a spectroradiometer.

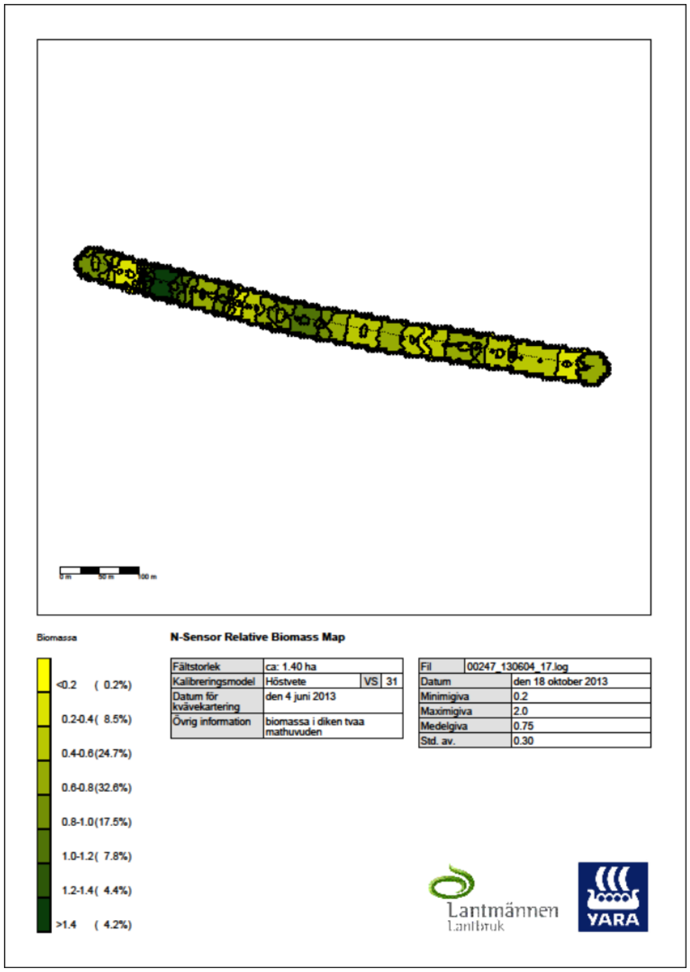



Figure 2.3: Output from a Yara N-sensor (Cederlund et al., 2014, pp. 48)

A portable field ***spectroradiometer*** such as the FieldSpec 3 Hi-Res (spectral range 350-2500 nm) or the FieldSpec 3 JR (spectral range 350-2500 nm) can be used for measuring plant reflectance. By using these measurements, the nitrogen and water levels in a plant can be monitored. Different targets reflect and absorb solar radiation differently, and because of this, they can be distinguished by their different spectral reflectance signaturesJensen (2007). One drawback of spectroradiometer instruments is that they often are very expensive.

**Infra-red spotlights** could be used in conjunction with **near infra-red**

**(NIR) cameras**. This would enable vision in (for humans) complete darkness, where the active infra-red light at one continuous frequency would be reflected by the environment. It is likely that the monitoring would be more accurate because of the stable and controllable input. When using the near-infra-red light emitted from the sun during daytime, the NIR frequencies vary depending on the time of day, time of year, and spatial location, including altitude.

The Danish National Railway *(Banedanmark)* took action against vegetation and implemented a system called WeedEye, which is a digital camera system for detecting weeds that was developed by Heisel and Christensen (1998). Problems with maintaining the system have meant that project has now been cancelled and is no longer used. Nonetheless, the paper is still quite interesting. The system has a sensor system for weed detection (leaf area of wheat and charlock) using digital cameras. Attached to each camera is a halogen bulb for enabling a continuous light spectrum. The camera system estimates the percentage of *visible leaf area*, which is based on information from the RGB colour space.

There is a lot of scientific literature about agricultural weed control, i.e., how to fight weeds. In the agricultural environment the following types of general approaches are usually used Perez et al. (1997):

1. Detecting certain geometric differences between the crop and weeds, such as leaf shape and plant structure.
2. Detecting differences in spectral reflectance, as in Philipp and Rath (2002).
3. Detecting differences in the location of the crops compared with the location of weeds.

Optical sensors have increasingly found more applications in agricultural production systems Wang et al. (2007). Advantages of this type of sensor over imaging systems include low cost, less-complex system configuration, and greater processing speed. Most optical sensors detect weeds based on plant spectral characteristics; The relation between wavelength and some other variable, e.g. between wavelength and emitted radiant power of a luminescent screen per unit wavelength interval, Alchanatis et al. (2005), Vrindts and De Baerdemaeker (1996) and Wartenberg and Schmidt (1999).

Another example of a commercial product is the Australian WeedSeeker (McIntosh Distribution, 2015). The WeedSeeker system works by using LEDs to project a combination of invisible infra-red and visible red light onto the target area approximately 60 cm below the sensor. A sensor captures reflected light, and analyses it, and then finds the light reflected by green plants. When a green plant's reflectance is identified, the sensor waits until the plant is under the spray nozzle and then triggers a fast-fire solenoid valve, which sprays the plant with herbicides. WeedSeeker is also used in the railway industry, as well as in agriculture.

## 2.3 Conclusions & Discussion

To stay competitive, both in terms of the budget and the environment, CBM is an important area to address. Using maintenance terminology for an automated system for the detection and characterisation of vegetation on railways would conform to the CBM paradigm, which includes condition monitoring. That is, system output will inform management personnel about the current state of vegetation based on the sensor information from the railway embankment and its surroundings, and/or make explicit suggestions about maintenance actions. Vegetation management, which is based on an analysis of the data collected during condition monitoring can directly and indirectly extend the life expectancy of railway objects; one of the most important of these is the railway embankment.

The approach taken in this work will, as suggested by Jardine et al. (2006), make use of the paradigms, techniques and methods from the field of artificial intelligence (AI) to identify faults when they occur and predict failures before they occur.

With regard to applicable sensors, many papers about agricultural weed control are of interest of this project, they can in some cases be mapped upon the railway embankment area. However, the environment of farming fields with their nutrient-rich soil in irrigated fields is quite different from the often harsh environment found on railway trackbeds and embankments.

# Chapter 3

# Plant Science

## 3.1 Motivation

Since the objects which are to be discovered, measured and characterised are *plants*, this thesis intersects with the area of biology and, more specifically, ecology and botany. For this reason an overview of selected fields of botany and ecology is justified and thus presented in this section.

## 3.2 Plant Requirements

***Question****: Why do some plants seem to flourish on railway embankments, which are meant to be an unfriendly and hostile environment?*

The ballast used on embankments will, over time, be ground down by high pressure into fine stone powder. In addition, it will inevitably become polluted with organic litter from dead plants, or oily and nitrogenous wastes from trains, leading to reduced support of the sleepers, reduced life of the wooden sleepers, and a softening and eventual failure of the sub-soil below Sargent (1984); (Trafikverket, 2012b, appendix 8); Persson, B., 8 June 2011, Personal interview. It has been suggested that this stone powder could be used, as is the case for granite powders, as a fertilizer for acid soils. Thus, this could give a positive effect Silva et al. (2005).

Since the sun warms up the ballast stones during the daytime, the stones will preserve the generated heat a long time after the sun has disappeared. The organic litter and stone powder, together with water, and the light and heat accumulated from the sun, will, like in any garden compost, be transformed into nutritious soil over time. All these factors will make some plants flourish in this environment.

The countermeasures used to stop this happening are to either clean or replace the ballast, or by managing the vegetation. The first option is an expensive process, which demands highly specialised machines, and as result of this, the replacement of ballast typically occurs at intervals of every 15th up to every 30th year on main tracks Profillidis (2006, p.338). Hence, for secondary tracks, the intervals might be even longer.

The second option is a short-term solution and is about the control of vegetation by means of mechanical, and/or chemical methods (using herbicides). Even so, at some point in time it will not help to use chemical or mechanical methods any more, because the ballast, or its surroundings, is too polluted with aggregated organic litter, and so forth (see above). Whilst these methods will stop plants from growing, they only bring about a temporary effect; a couple of weeks later, vegetation will flourish once again. The absence of vegetation control means that all the necessary basic plant environmental requirements have been fulfilled over time. This is the time to renew or replace the ballast.

## 3.3 Analysing and Classifying Plants

In order to represent plants and be able to efficiently communicate about them in an structured manner, it is indeed relevant to give a brief review of some scientific classifications in biology.

One way of classifying plants is to make use of various systematic approaches to *botanical classification*. These are usually based on the Swedish taxonomy introduced by Carolus Linnaeus (1707 - 1778), often called the Linnaean System Campbell and Wynne (2011).

All known living organisms are arranged in an hierarchical structure. One of the latest commonly accepted *botanic taxonomic hierarchies* has six top ranks called *kingdoms* Cavalier-Smith (2004). One of these six kingdoms relates to plants, and, at the bottom of such an hierarchy, can be found the most specialised rank, known as species (see table D.1 in D).

Another approach is to analyse the *physiognomy* and the *structure of plants* Kuchler (1966). Physiognomy refers to the physical (life) forms of plants (see the left-hand side of table 3.1).

With regard to structure, Kuchler focuses on the two characteristics of *height* and *coverage* and, while it is often hard to measure them exactly, he suggests a form of vegetation mapping, as seen in table 3.1 (right-hand side).

| Life Form Categories | | |
|---|---|---|
| **Basic Life Forms** | | **Special Life Forms** |
| ***Woody Plants*** | | Climbers |
| Broadleaf evergreen | B | Stem succulents |
| Broadleaf deciduous | D | Tuft plants |
| Needleleaf evergreen | E | Bamboos |
| Needleleaf deciduous (e.g Larix) | N | Epiphytes |
| Aphyllous | O | |
| Semideciduous (B + D) | S | |
| Mixed (D + E) | M | |
| | | |
| *Herbaceous plants* | | |
| Graminoids | G | |
| Forbs | H | |
| Lichens and mosses | L | |

| Structural Categories | | | | |
|---|---|---|---|---|
| **Height** | | **Coverage** | | |
| *Class* | *Height* | | Class | Coverage |
| 8 | > 35m | | c | continuous (>75%) |
| 7 | 20-35m | | i | interrupted (50-75%) |
| 6 | 10-20m | | p | parklike, in patches (25-50%) |
| 5 | 5-10m | | r | rare (6-25%) |
| 4 | 2-5m | | b | barely present, sporadic (1-5%) |
| 3 | 0.5-2m | | a | almost absent, extremely scarce (<1%) |
| 2 | 0.1-0.5m | | | |
| 1 | < 0.1m | | | |

Table 3.1: A method of analysing plants by their physiognomy and structure devised by Kuchler (1966)

## 3.4 Biophysical Data, and Root and Shoot Ratios

The term *biophysical* data refers to measurements of physical characteristics collected in the field. Examples of such data are type, size, biomass weight, spacing between plants, form, texture and/or mineralogy of soil Campbell and Wynne (2011, p.384). When it comes to weight, biomass root to shoot ratios are often used. A root to shoot ratio describes the ratio between the shoot (i.e., the above-ground part of the plant) and the root, and is one of several studied areas in *allometry*.

In order to model plant growth Pearsall (1927) applied the formula in equation 3.1

$$y = b * x^{\alpha} \tag{3.1}$$

where $y$ is the dry weight of the root, $x$ is the dry weight of the shoot, and $b$ and $\alpha$ are constants.

Allometry is the study of biological scaling. The term in it self was (probably) first mentioned by Huxley and Teissier (1936), who recognised that many scaling relationships, when plotted on a log-log scale, were linear. Consequently these relationships could be described (by rewriting equation 3.1) using the simple linear equation 3.2, as follows:

$$log\,(w) = log\,(b) + \alpha * log\,(z) \tag{3.2}$$

where $z$ is body size, $w$ is organ size, log $b$ is the intercept of the line on the $w$-axis and $\alpha$ is the slope of the line, also known as the allometric coefficient, Shingleton (2010).

The equation can also be applied to investigating the relative distribution of growth of shoot to root Ledig et al. (1970) as follows:

$$log\,(shoot\,dry\,weight) = log\,(b) + \alpha * log\,(root\,dry\,weight) \tag{3.3}$$

The relationship between shoot biomass and root biomass (often denominated $M_A$ and $M_B$, respectively) is often used to estimate a growth response compared with surrounding environmental conditions, or to evaluate the responses of individual plants to experimental manipulation, e.g. Monk (1966), Hunt and Lloyd (1987) and Niklas (2005).

It has been shown that, in general, one could roughly approximate the root to shoot ratio of Norway spruce and Scots pine as one to two year old seedlings as: $R = 0.2 * S$ where $R$ is the dry weight of the root, and $S$ is the dry weight of the shoot. It is also commonly observed that one year old nursery generated Scots pine and Norway spruce seedlings have a root-shoot ratio of 0.25, i.e. of the seedlings total dry weight biomass, the dry weight of the root is approximately 20% and the shoot 80%, respectively (Lindström, A., 7 Sept. 2011, *Personal interview*).

For more information, a thorough review about modelling the root to shoot ratios has been made by Wilson and Bastow (1988).

## 3.5 Common Woody Plants along Railways

**Motivation:** One of the goals was to recognise those woody plants that are most likely to grow on railway embankments. The woody plants and especially conifers are hard to kill by the use of legal herbicides. By knowing the where these kind of plants are in the railway system, it might be feasible to mechanically harvest them.

Thus, a short summary about the most common species in this section is justified here, in the sense that a physician can benefit from knowing his patients. The analogy applies in this work, but this time, it concerns plants.

Woody plants, also known as ligneous, or wood-like plants, refer to trees and bushes that have a wood-like stem and branches.

Forest statistics by Swedish National Forest Inventory 2007-2011 and 2006-2010 SFNI (2011) reported that the three most common kinds of trees in Sweden are *Picea abies* (Norway Spruce) followed by *Pinus sylvestris* (Scots Pine), followed by *Betula* (birch).

**Picea abies (Norway Spruce)** is Sweden's most common species and is found throughout the country. The spruce thrives in fertile and slightly moist soil and prefers shade. In southern Sweden, a southern variant with wide crown and horizontally outward branches dominates, whilst in the north, a variant with narrow crown and downward branches prevails. The spruce has a shallow root system and is therefore sensitive to the wind.

The Norway spruce is a *secondary tree species* meaning that it prefers to establish itself where other primary tree species are already established.

Primarily, this is because the spruce prefers shady environments. After seeds have settled, it is common that a one-year old naturally generated Norway spruce seedling usually grows slower than a Scots pine seedling during the first two years, which means that it will be a little shorter than the Scots pine in height (Lindström, A., 7 Sept. 2011, *Personal interview).*

Figure 3.1: Picea abies seedling established on sleepers

**Pinus sylvestris (Scots Pine)** grows throughout Sweden, like the spruce, but is most common in northern Sweden as well as on Gotland. The Scots pine prefers light, and grows preferably in dry soils and on peatlands. The pine has a pole-like root which makes it more wind resistant than the Norway spruce.

The Scots pine is an invasive primary tree species, meaning that it prefers to propagate by invading open and light areas (Lindström, A., 7 Sept. 2011, *Personal interview)*.

A one-year old naturally generated Scots pine seedlings could grow to between 2 and 7 cm in height. During the second year it is normal for the seedling to produce an apical shoot of approximately 10 cm. Similarly, a one-year old Scots pine seedling generated in a nursery can grow to a height of between 10 and 15 cm (Lindström, A., 7 Sept. 2011, *Personal interview)*

Figure 3.2: Pinus sylvestris seedling established on a sleeper

A fully grown tree is usually 20 to 30 metres in height and can be as old as 900 years. Pine needles are arranged in pairs and are blueish-green to greyish-green in colour. The needles are 3 to 6 cm long (Nationalencyklopedin, 2011).

The genus **Betula (birch)** is widespread throughout the Northern Hemisphere, across a number of different climate types, including subarctic, mountainous and temperate zones. There are two Betula species in Sweden, namely Betula pubescens (downy birch) and Betula pubescens ssp. czerepanovii (mountain birch, or Arctic downy birch), which is a variant of the downy birch and Betula pendula (silver birch). The birch is the most common leaf tree is found growing throughout the whole of Sweden. Downy birch is most prevalent in northern Sweden, where up to 75% can be found.

Figure 3.3: Betula seedling establishment on ballast

After the spruce and pine, the birch is the most common tree species. Both the birch and the Scots pine are invasive primary tree species, meaning that they prefer to invade open and light areas.

### 3.5.1 Woody Plant Candidates for Data Collection

In addition to the species mentioned above, this work also includes the species listed in table 3.2. Their inclusion is dependent on standing volume statistics reported by SFNI (2011) and suggestions made by domain experts (Lindström, A., 7 Sept. 2011, *Personal interview*) and (Stattin, E:, 28 June 2011. *Personal interview*). On the advice of these, some species have been excluded, even though their ranking in the standing volume statistics is quite high; for example, Fagus sylvatica L. (beech). Similarly species that are likely to spread out on embankments have been included, even though they rank lower, or are not present in the standing volume statistics, e.g. Acer platanoides (Norway maple), and Prunus padus (bird cherry).

| **Familia** | **Genus** | **Species** | **English Name** |
|---|---|---|---|
| *Pinaceae* | *Picea* | *Picea abies* | *Norway Spruce* |
| *Pinaceae* | *Pinus* | *Pinus sylvestris* | *Scots Pine* |
| *Betulaceae* | *Betula* | *Betula pubescens* | *Downy birch* |
| *Betulaceae* | *Betula* | *Betula pendula* | *Silver birch* |
| *Betulaceae* | *Alnus Mill.* | *Alnus incana* | *Grey Alder* |
| *Betulaceae* | *Alnus Mill.* | *Alnus glutinosa* | *Alder* |
| *Salicaceae* | *Populus* | *Populus tremula* | *Aspen* |
| *Rosaceae* | *Sorbus* | *Sorbus aucuparia* | *Rowan* |
| *Fagaceae* | *Quercus* | *Quercus robur Pedunculate* | *Oak* |
| *Salicaceae* | *Salix* | *Salix caprea* | *Goat Willow* |
| *Rosaceae* | *Prunus* | *Prunus padus* | *Bird Cherry* |
| *Aceraceae* | *Acer* | *Acer platanoides* | *Norway Maple* |

Table 3.2: Plants likely to be found on railway embankments

# Chapter 4

# Measuring Terrestrial Vegetation

In this chapter, the process of measuring, or quantifying vegetation is described. One way of quantifying a population of interest is obviously to count, or measure the full population. Often, it is not feasible to do so because of three types of constraints (Barnett, 2003): 1) time, 2) money, and 3) accessibility. Another commonly used approach is to study a smaller group, i.e., a *sample* of the total population.

## 4.1 Sampling Units

In general, a *sampling unit* is an element or set of elements that is considered for selection out of the total population. In the context of this work, the most often used primary sampling units are individual plants, plant clusters, lines (transects), points, plots often denoted quadrats. Combinations can be used, such as point quadrats (i.e., points within a plot), sub plots (i.e., small plots, which can be composed of a grid within a bigger plot, such as the frequency method), and plots sampled along a line. The type of sampling unit will be dependent on the type of vegetation attribute being measured Elzinga et al. (1998, p.101).

## 4.2 Spatial Plant Distribution

With regard to plants, there are three types of small-scale basic spatial patterns, or geographic local distributions Krebs (1999); Mauseth (1998):

1. Random distributions (see figure 4.1 a) having unpredictable spacings
2. Aggregated/clumped distributions (see figure 4.1b)
3. Uniform distributions (see figure 4.1 c)

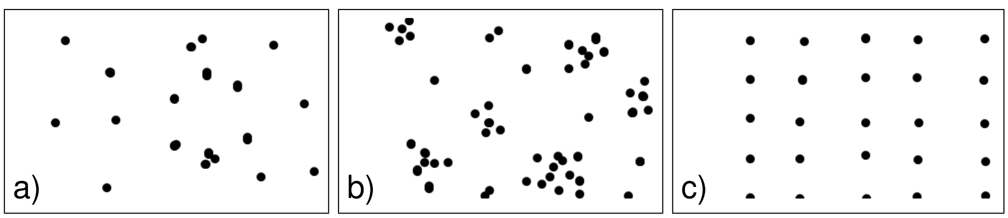



Figure 4.1: a) Random, and b) aggregated, and c) uniform plant patterns

These three types of spatial distributions (also called *dispersions*) are the spatial relationships of individual organisms to one another. It is important to know the type of spatial pattern in order to describe and predict plant populations. If the spatial distribution is accounted for, the sampling strategies can increase sampling efficiency Cardina et al. (1997) and Cardina et al. (1995).

Krebs (1999, pp. 115 - 123) describes how to investigate whether *randomness* apply for a spatial pattern, the Poisson distribution is the most common descriptor of data. Randomness in two dimensions is here defined as randomised x- and y-coordinates in a geographical space. The Poisson distribution is described as in equation 4.1.

$$P_x = e^{-\lambda} \left( \frac{\lambda^x}{x!} \right) \tag{4.1}$$

where $P$ is the probability of observing (by counting) $x$ number of individuals in the sample area, and $\lambda$ is the true mean of the distribution. To fit this distribution to observed data the investigator only has to estimate the mean by setting $\overline{x} = \lambda$, where $\overline{x}$ is the sample mean. Next step is to test if the Poisson distribution provides an acceptable fit to the sampled data set. For this distribution there are in general two tests for the goodness of fit: the *index of dispersion test* and the *chi-squared goodness of fit test* (see equation 4.2). The latter one is less sensitive than the index of dispersion test.

$$\chi^2 = \sum \left( \frac{(O - E)^2}{E} \right) \tag{4.2}$$

where $O$ is the observed frequency (i.e. the number of individuals in each sample area), and $E$ is the expected (theoretical) frequency. The null hypothesis $H_0$ can be formulated as: the data are consistent with the Poisson distribution (i.e. there is no difference between the observed and the expected distribution). The alternative hypothesis states the opposite, i.e. the data is not consistent with the expected values for the specified distribution. If the observed $\chi^2$ value is larger than the tabulated value (for the chosen significance level $\alpha$, often 0.05), then $H_0$ is rejected. It follows that if $H_0$ is *not* rejected then spatial randomness is assumed.

In order to investigate if the spatial pattern is *aggregated* (aka clumped, clustered) then Krebs (1999, p. 124) suggests the *negative binomial distribution* is the most common one. Note that this distribution cannot act as a descriptor for all aggregated patterns, as there are an infinite number of them. The *negative binomial distribution* could be described as; the probability of the event $X$, where $X$ is: *Observing $x$ number of individuals in one sample area* (see equations 4.3, 4.4, 4.5, and 4.6 as examples for finding 0 to 3 individuals).

$$P(X = 0) = \left(1 + \frac{\overline{x}}{k}\right)^{-k} \tag{4.3}$$

$$P(X = 1) = \left(\frac{k}{1}\right) \left(\frac{\overline{x}}{\overline{x} + k}\right)^{1} \left(1 + \frac{\overline{x}}{k}\right)^{-k} \tag{4.4}$$

$$P(X = 2) = \left(\frac{k}{1}\right) \left(\frac{k+1}{2}\right) \left(\frac{\overline{x}}{\overline{x} + k}\right)^{2} \left(1 + \frac{\overline{x}}{k}\right)^{-k} \tag{4.5}$$

$$P(X=3) = \left(\frac{k}{1}\right)\left(\frac{k+1}{2}\right)\left(\frac{k+2}{3}\right)\left(\frac{\overline{x}}{\overline{x}+k}\right)^3\left(1+\frac{\overline{x}}{k}\right)^{-k} \tag{4.6}$$

where $P(X = x)$ is the probability of finding $x$ number of individuals in a sample area (aka quadrat); $\mu$ is the mean; $k$ is the negative binomial exponent. When trying to fit the negative distribution with some real world data there are two parameters to be estimated; the mean $\mu$ and the negative binomial exponent $k$. The mean is set to the observed sample mean as: $\overline{x} = \mu$. The negative binomial exponent $k$ is harder to estimate. It is suggested by Anscombe (1950) in Krebs (1999, p. 125 - 129) to first approximate $k$ (denoted as $\hat{k}$) as in equation 4.7:

$$\hat{k} = \frac{\overline{x}^2}{s^2 - \overline{x}} \tag{4.7}$$

where $s^2$ is the sample variation and $\overline{x}$ is the sample mean. This approximation of $k$ works well if the number of sample areas (i.e. quadrats) are large i.e. are more than or equal to 20 and the counts cannot be arranged in a smooth distribution. If the number of sample areas, i.e. quadrats, is less than 20 (probably not very likely when sampling on railway track sections), then see the flow-chart accompanied with suggested methods in appendix J.

Next, three tests are appropriate for testing if the negative distribution can describe the observed data set. Those are: Chi-squared Goodness-of-fit (see equation 4.2), U-statistic Goodness-of-fit , and T-statistic Goodness-of-fit. Especially when the sample size is less than 50, the last two tests are better to detect departures from the theoretical negative binomial distribution than the chi-squared test Krebs (1999, p. 135).

It has been shown that local plant populations in general are not randomly distributed; rather, they are aggregated (clumped) Mauseth (1998) i.e., the distances between plants are either small or large, but seldom are they average or uniform Cardina et al. (1995) and Cardina et al. (1997).

### Index of dispersion for sample area counts

An *index of dispersion* is a *normalised measure* of the dispersion of a probability distribution. It is used to quantify whether observed data (e.g.the counting of plants) have spatial aggregated pattern, a random pattern, or a uniform pattern compared to a standard statistical model, for example the Poisson distribution. The special thing about such an index is that it should not ideally be affected by sample size, population density, or by variation in the size and shape of the sampling quadrat Krebs (1999, p. 212).

When choosing which method to use to *determine the degree of dispersion*, the standardised Morisita Index Morisita (1959) and Morisita-Horn index (Horn, 1966) are seen to be among some of the better methods Wolda (1981). This is because these methods are nearly independent of sample size Wolda (1981) and Krebs (1999).

Krebs (1999, p. 222) points out that the spatial pattern obtained, and the resulting index of dispersion is dependent on the size and shape of the quadrat. In figure 4.2 there are 16 + 9 + 4 + 1 = 30 sub-plots in total. Dependent on the chosen size of the quadrat (and of course the sampling scheme) different results may come up. The only solution to this problem is to sample the population with different quadrat sizes and record how the index of dispersion changes with the size.

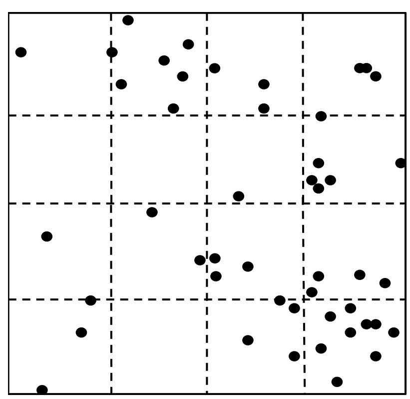


Figure 4.2: Example of distribution dependent on scale of sample square

## 4.3 Vegetation Attributes

Plant species can be described by a number of quantitative features or characteristics called *vegetation attributes*. These describe how much, how many, or what kind of plant species are present. In general, the most commonly used attributes when monitoring are: *cover*, *density*, *frequency* and *biomass,* which is a measure of *production* Elzinga et al. (1998, p.101) and Bonham (1989).

Several approaches are used to estimate, or measure cover, depending on the sampling unit, but in general they can be classified as: lines, points and quadrats (see section 4.1).

### 4.3.1 Plant Cover

In ecology, plant cover is one of the most commonly used variables for monitoring ground state Bonham and Clark (2005) and (Jukola-Sulonen and Salemaa, 1985).

Usually, cover is defined as the vertical projection of vegetation from the ground, as viewed from above, i.e., a bird's-eye view of the vegetation. The attribute cover is usually expressed as a percentage (see equation 4.8).

$$Cover = \frac{Ground\,surface\,area\,covered\,by\,vegetation}{Total\,ground\,surface\,area} * 100 \tag{4.8}$$

The vegetation in equation 4.8 could also be replaced by other coverages of interest, such as rocks, litter, or bareground etc.

Elzinga et al. (1998, p.178), described two types of cover (see figure 4.3): 1) Basal cover defines the area in which the plant intersects the ground, and 2) Aerial cover is the vegetation that covers the surface above the ground. With regard to aerial cover, two types can be distinguished, namely aerial foliar cover (AFC) and aerial canopy cover (ACC) *Coulloudon et al. (1999, p.25)* (see figure 4.4).

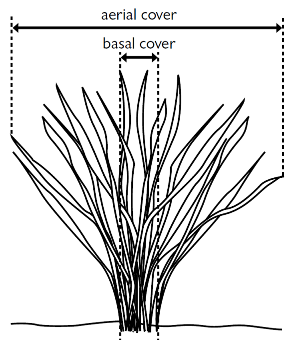



Figure 4.3: Aerial cover vs. Basal cover Elzinga et al. (1998, p.178)

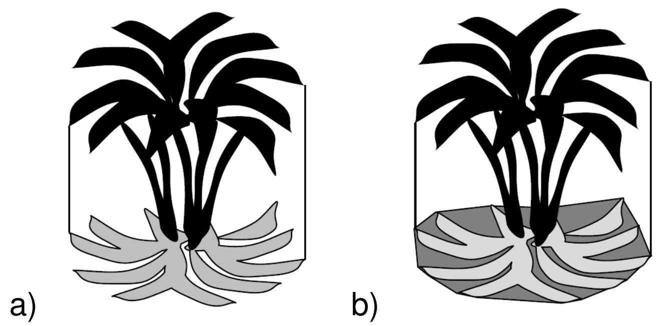



Figure 4.4: a) Aerial foliar cover vs. b) Aerial canopy cover Coulloudon et al. (1999, p.25)

Foliar cover and canopy cover have been defined as:

*Aerial foliar cover (AFC)* is the area of ground covered by the vertical projection of the aerial portions of the plants. Small openings in the canopy and intra-specific overlap are excluded *(see figure* 4.4a).

*Aerial canopy cover (ACC)* is the area of ground covered by the vertical projection of the outermost perimeter of the natural spread of foliage of plants, also known as the convex hull. Small openings within the canopy are included (see figure 4.4b). If more than one species is to be included in the total cover, the canopy cover may exceed 100% because of overlapping Coulloudon et al. (1999, p.25).

The attribute cover *is not biased by the size and distribution of individuals* and can therefore be used to compare the abundances of species of widely different growth forms (Whittaker, 1975); (Floyd and Anderson, 1987). Cover is the attribute which is most directly related to biomass out of the three attributes: density, frequency, and cover Elzinga et al. (1998).

It is important to observe that cover changes during a growing season and, therefore, sampling must be done at the same time each year. In addition, the current year's weather history also has a great impact on cover Elzinga et al. (1998, p.179).

Depending on the author, different types of cover go under different names, and are also interpreted differently. This partly depends on *how cover is defined*; that is to say, how and what is measured when estimating cover. Fehmi (2010) compared published common plant cover definitions as well as uses of cover in research published between 1950 and 2007. In order not to limit the survey, three overall definitions were made in an attempt to incorporate all the authors' cover definitions. The three suggested overall cover definitions identified while conducting this comparative survey were: 1) Aerial cover, 2) Species cover, and 3) Leaf cover (see figure 4.5). As seen in the figure, the total cover percentage result is dependent on how the rater chooses to define cover.

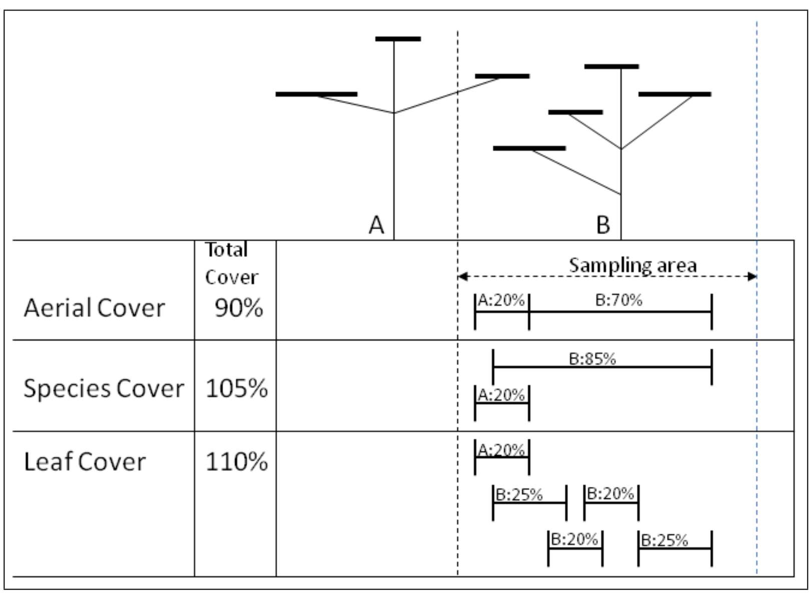



Figure 4.5: Three definitions of cover

The most common approaches used to measuring cover are:

- visual estimates VE in plots, or
- line interception, or
- point interception, or
- sub-plot frequency

For a detailed analysis in which these methods are compared, see for example Hurford (2006), Mueller-Dombois and Ellenberg (1974), Brakenhielm and Qinghong (1995), Bonham (1989) and Jonasson (1988), where the authors considered the point interception approach to be the least biased and most objective of the three basic cover measures. In order to calculate the accuracy of compared cover-measuring methods, manual image processing was used to measure the "true" value i.e., percentage coverage in images. Brakenhielm and Qinghong (1995) measured the cover percentage of two completely visible species from images (Vaccinium myrtillus (Bilberry), and Vaccinium vitis-idaea L. (Cowberry, or Lingonberry)). Firstly, images were scanned into a computer. A drawing was then made manually to outline the visible parts belonging to the same species. In conjunction with this manual operation, automatic outlining based was applied by thresholding the brightness of the leaves to separate leaves from gaps.

### Plots

When measuring cover using plots the most common method is to make a visual estimate and map it to a cover class, i.e., by mapping the plot estimation to, say, 35% to an interval (e.g. Daubenmire class 3). However, as cover is visually estimated, variations between estimated samples can occur, especially if more than one person surveys the vegetation. For this reason, cover percentages are normally converted into cover classes, which are placed on a scale of a certain number of intervals between 0 to 100%. The various cover class systems help to compress errors. Visual cover estimates that use classes such as these (i.e., coarse grade scales) are usually reliable enough for categorizing different types of vegetation communities (Mueller-Dombois and Ellenberg, 1974).

With regard to estimating cover in plots, there are several cover class systems from which to choose. According to Elzinga et al. (1998), among them, the most often used include the *Braun-Blanquet* Braun-Blanquet (1932), and the *Daubenmire* Daubenmire (1959) systems. In the UK, a detailed phytosociological classification called the National Vegetation Classification (NVC) was carried out using the *Domin-Krajina* system rather than the two mentioned above. This was confirmed by Hill (2005, p.203) who also stated that the Domin scale is the most used, and credited its use in the NVC as one example. The classes of the three systems can be viewed in table 4.1.

| **Braun-Blanquet** | | | **Domin-Krajina** | | | **Daubenmire** | | |
|---|---|---|---|---|---|---|---|---|
| ***Class*** | ***Cover (%)*** | ***Mean*** | ***Class*** | ***Cover (%)*** | ***Mean*** | ***Class*** | ***Cover (%)*** | ***Mean*** |
| 5 | 75-100 | 87.5 | 10 | 100 | 100.0 | 6 | 95-100 | 97.5 |
| 4 | 50-75 | 62.5 | 9 | 75-99 | 87.0 | 5 | 75-95 | 85.0 |
| 3 | 25-50 | 37.5 | 8 | 50-75 | 62.5 | 4 | 50-75 | 62.5 |
| 2 | 5-25 | 15.0 | 7 | 33-50 | 41.5 | 3 | 25-50 | 37.5 |
| 1 | 1-5 | 2.5 | 6 | 25-33 | 29.0 | 2 | 5-25 | 15.0 |
| *†* | <1 | 0.1 | 5 | 10-25 | 17.5 | 1 | 0-5 | 2.5 |
| *r* | <<1 | * | 4 | 5-10 | 7.5 | | | |
| | | | 3 | 1-5 | 2.5 | | | |
| | | | 2 | <1 | 0.5 | | | |
| | | | 1 | <<1 | * | | | |
| | | | *†* | <<<1 | * | | | |

Table 4.1: Plant cover class systems (Mueller-Dombois and Ellenberg, 1974)

## Point Interception Methods

The basic outline of point interception methods is that the vegetation gets vertically intercepted point-wise with, for example, steel, wooden pins, or a laser, or by looking through a cross-hair. Each time the vegetation under observation is physically hit by a pin, a counter is increased by one. The cross-hairs in a grid (e.g. a frame grid made of cord or steel wires) can also be regarded as a set of pins, where each cross-hair serves as a pin. The cover percentage for a species in each layer of the canopy is simply the number of hits out of 100, or any other integer. Several techniques can be applied as follows:

*Single Point Interception (SPI)*: Each sample point is defined by a sampling pole positioned vertically to the ground. Often used along transects where many SPIs are systemically deployed (i.e., sampled). It has been shown that SPI can give more precise cover estimates than points grouped into point frames or grid frames, given that the same number of points are sampled Goodall (1952) and Greig-Smith (1983). Using individual points requires approximately 1/3 of the number of points used when in groups, such as in PQI and PFI Bonham (1989, p. 110).

*Point Quadrat Interception (PQI)*: Each time an individual species is observed under an intersection in the quadrat grid (i.e., a cross-hair), it is counted as a hit. Figure 5.4 shows a point quadrat grid (single layer). This grid frame has 81 cross-hairs, and consequently 100 sub-plot squares. In this context, the cross-hairs are used to intercept the vegetation. A hit equals 1/81 (~1.2%) total cover, and, if desired, the total cover for each species can also be calculated. This method is often used to estimate cover of short vegetation (<20 cm) Bonham (1989, p. 110). The method is precise and is one of the most objective measuring methods; however, it is also time consuming and, in addition, needs to be conducted by experts Hill (2005, p.217). The same grid frame can be used for the sub-plot frequency method. In this case, the sub-plots are of interest, not the intersections (see section 4.3.2)

*Point Frame Interception (PFI):* Each time an individual species is observed, i.e., hit by one of the 10 pins (as in figure 4.6), it is counted as a hit. The number of pins in the frame can vary, and this influences the efficiency of sampling.

Figure 4.6: 10-pin point frame (Bonham, 1989)

No matter which of the techniques is used, the distance between the pins (and the distance between each point of sampling), for example along a transect, depends on: 1) plant size, 2) plant distribution, and 3) the distance between plants.

### 4.3.2 Plant Frequency

The frequency attribute is a measure of presence/absence (also hit/miss) in a predefined area. Either the plant in question is present in the predefined area, or it is absent. It is a boolean variable. In general, the total number of area hits divided by the total number of examined areas is denoted the frequency rate Mannetje and Jones (2000, p.86) and Hill (p.11 2005). Frequency is measured using quadrats and is affected by the quadrat size; the latter may mean that it is less meaningful than other measurements. If a quadrat is too small, then rare plants may not be recorded. On the other hand, if a quadrat is too large, then individual species are likely to occur in all quadrats and frequency values will equal 100%. This will not allow increases or decreases in frequency to be monitored.

A benefit of using frequency as a method of measurement is that it is less tedious and less time consuming than measuring density, cover and production attributes such as biomass.

### Frequency in Quadrats

Quadrats of different shapes and sizes are used to measure frequency. Frequency is the percentage of total quadrats that contain at least one example of a given species (see equation 4.9.

$$Frequency = \frac{(No.\, of\, quadrats\, in\, which\, species\, occur)}{\sum(Quadrats\, sampled)} \qquad (4.9)$$

*Relative frequency* of one species as a percentage of total plant frequency, see equation 4.10.

$$Rel.\, Frequency = \frac{(Species\, frequency)}{\sum(frequency\, values\, for\, all\, species)} * 100 \quad (4.10)$$

### Sub-plot Frequency

An extension if the frequency measure is the *sub-plot frequency*, where the measurement of interest is the number of sub-plots that contain the target species Hill (p.11 2005). In figure 4.7 22 target plants are represented as dots. If the thick frame had been the sampling unit, it would have been counted as 1. If the grid had been used instead (i.e., 10-by-10 sub-plots), then the count would have been 13, because the target plants are present in 13 of the 100 sub-plots (and absent from 87). If the sampling frames (i.e., the thick frame and the grid of 100 sub-plots) had been used (e.g. randomly placing the sampling unit over a larger field), then the sub-plot version would have yielded a more finely granulated frequency result. The same grid sample frame can be used for estimating cover by using the point interception method (see section 4.3.1).

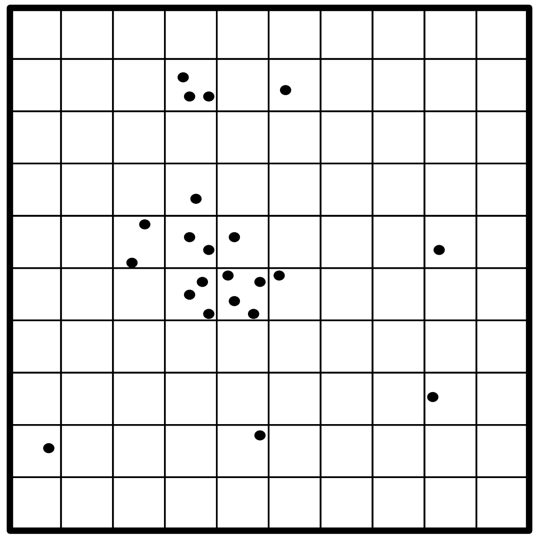


Figure 4.7: Estimating frequency

### 4.3.3 Plant Density

The density attribute (more precisely known as absolute density) is a measure of abundance. The density is defined as the number of individuals per unit area or volumeKrebs (1999), Elzinga et al. (1998) and Hill (2005).

This thesis only deals with area units, as defined in equation 4.11. Thus, it is all about counting individuals in a predefined area, for example a Picea abies (Norwegian spruce) density of ten Picea abies per square meter.

$$Density = \frac{No.\ of\ individuals}{Area\, sampled} \tag{4.11}$$

as defined in Eq.4.12

$$Relative\, Density = \frac{species\, density}{total\, density\, for\, all\, species} * 100 \tag{4.12}$$

A drawback with using density estimates is that they require an individual plant to be defined. This may be difficult sometimes, especially if flowering herbs are not in blossom. For example, even professional botanists find it difficult to determine which species plants from the common familia Asteraceae (Swe: Fibblor) belong to (Stattin, E., 28 June 2011. *Personal Interview*).

### 4.3.4 Plant Biomass: A Measure of Production

Estimating *plant biomass* (phytomass) is a central part of many ecological investigations Jonasson (1988); Brakenhielm and Qinghong (1995). It is used for characterising ecosystems, or, for example, measuring productivity (production), and vigour. Biomass falls under the vegetation attribute production. Destructive methods of harvesting are often used to extract biomass, meaning that the root is excavated. Sometimes, the part of the plant found above ground (often called the shoot) is also excavated, depending of the focus of the study. The next step is often to weigh the excavated plant biomass. Both fresh weight and dry weight can be used. If the fresh weight is of interest then the weighing process must be done immediately after the excavation, otherwise the plant dies or is deformed because of lack of water, light, nutrients, etc. (see section 3.2). The fresh weight can vary substantially even for the same plant because of, for example, temporary changes in weather. A healthy plant that soaks up as much water as it demands will weigh more than the same plant during/after dry periods. A way to avoid this issue is to measure the dry weight. After this, all minerals, (i.e., the soil) has to be taken away. If the root is about to be separated from the shoot, then it is important to decide where to make the cut. Depending on the plant species, this can have an effect on the ratio between the root and the shoot, because of the small weights involved (often down to milligrams).

## 4.4 Conclusions & Discussion

The most common approaches for manually measuring cover are to use visual estimates (VEs) in plots, line interception, point interception, and sub-plot frequency. For the transfer to machine vision, this way of measuring plant cover and frequency seems to be most useful. In addition, since a raster image is a grid in itself, where each square at the lowest level is a pixel, the sub-plot frequencies were found to be both interesting and useful. This is because a raster image in itself is a very fine granulated grid, depending on the resolution. For example an image size of 800 x 600 pixels compromises 480 000 potential sub-plots.

A problem with the density attribute is that it requires individual plants to be recognised. Still, by redefining the protocol for what to count when measuring density, this attribute could be useful. For example, if a definition of plant clusters can be obtained, then a modified density attribute could involve both clusters and individuals. With regard to production attributes, measuring plant heights was found to be highly applicable. This could be carried out using optical laser measurements from above.

To be able to describe and predict plant populations on railway trackbeds /embankments the type of spatial pattern must be determined. In general there exists three types of spatial patterns: the random, the uniform, or the aggregated (aka clustered, or clumped) pattern. To determine this the Poisson and the negative binomial distribution could be used as tools to decide whether the spatial pattern is random or aggregated respectively.

To determine the degree of dispersion using an index of dispersion, the Standardised Morisita Index and the Morisita-Horn index are recommended, since these methods are nearly independent of sample size.

Much of the existing literature on agricultural weed control may relate to this project. In some cases, such research can be mapped onto the railway embankment environment. It should be noted that farm fields with their nutrient-rich and irrigated soil are quite different to the harsh environment found on railway embankments.

# Part III

# Data Collection

# Chapter 5

# Field Measurements and Experiments

All field data in this thesis were acquired on operational, lightly trafficked railway tracks, usually frequented by goods trains. Choice of lightly trafficked railway sections was motivated by the fact that such tracks possessed realistic amounts of vegetation. In addition, low traffic levels enabled us to work efficiently, which would not have been the case on main lines, where a train might pass every ten minutes. The nature of collecting data in such an environment meant that various safety and security inspectors were hired and present at all times.

### Image Data

Any image unloaded from a camera in use is represented as a Red-Green-Blue (RGB) image in RAW format. Two general types of images used in this thesis are:

1. VIS images, i.e., RGB images that represent the human visual spectrum (approx. 400 to 700 nm). The colours of an object in an image appear to a human observer the same way as if this observer were to directly view the object.

2. Grey scale images. Each image pixel is stored as a byte, a grayscale value that can be represented in $2^8 = 256$ shades of grey in between the extremes 0 (for black) up to 255 (for white).

Each image was tagged with a geographical position that shows the photo point, i.e., the camera's spatial location in longitude, and latitude.

## 5.1 Overview

This chapter describes the field experiments carried out in order to acquire data from railway trackbeds/embankments. As described in section 1.2, different inspectors may, for various reasons, interpret the regulations in different ways or have different opinions of what or how to inspect. Thus, it is difficult or even impossible to objectively, uniformly and consistently describe the state of the vegetation growing along railway lines. Thus, it was thought desirable to automate the manual inspection procedure. To adopt an automated procedure, *quantitative measurements and well described systematics had to be used.*

### 5.1.1 Camera Sensors

As mentioned in section 1.3, the sensors needed to be available on the commercial market at a reasonable price so that most subcontractors would be able to afford them. During the field experiments, a DSLR Nikon D90 camera and a DSLR Nikon D200 camera were used to sense the visual spectrum (approximately 400 to 700 nm). Extra illumination, or flash were not used. The output images were available in RAW format as well as JPG files in the RGB colour space. Image sizes were as follows: Nikon D200 2096*1944 pixels resolution, and Nikon D90 4288*2448 pixels resolution, respectively.

Identical lenses were used for both cameras: Tamron Ultra wide-angle lens SP AF10-24mm F/3.5-4.5 Di II LD Aspherical [IF], enabling an angle of view up to 108°44' at 10 mm focal length.

The acquisition of data was carried out in two ways: first, by recording the results of manual sampling from sample areas, as described in (Elzinga et al., 1998), (Coulloudon et al., 1999) and (Barker, 2001), and second, by acquiring images of the same sample areas.

All images were acquired outdoors, and were not illuminated by artificial lighting. The documentation of lighting conditions was conducted by reporting subsets of weather observations. These followed chosen subsets of the World Meteorological Organization's of the SYNOP protocol, which is basically a code system for reporting surface observations from land. (Unisys Weather, 2012).

### 5.1.2 Information Storage

After exporting the image files from each camera they were stored in a file system on a backed-up network drive. Coupled with every image was so-called Exif information, which is essentially meta-data about each image and how the camera was set up for any given shot. Examples of stored meta-data information are date and time when the image was acquired, shutter speed, focal length, and exposure details. Exif information is stored along with all data about the sensors. A database model was designed to be able to manage all the relevant files (see appendix F). A database was implemented based on the database model.

## 5.2 Motivation for the Field Experiments

A target population was defined using a universal list that classified plants by their physiognomy and structure (see table 3.1 in section 3.3) of this thesis. The target population included all woody plants, forbs, graminoids, and lichens and mosses (see table 3.1) found growing on railway embankments in Sweden. The target population covers a huge area of Sweden, through which approximately 14000 kilometres of railway runs. In order to narrow the scope some constraints had to be applied. The requirements for plant growth (see section Section 3.2), were selected as attributes to be able to control the scope of the population. Certain plants grow in certain environments, depending on the magnitude of each of the plant growth requirements. This is partially monitored by the Swedish National Forest Inventory (SNFI) (Swedish University of Agricultural Sciences, 2011) and presented in their statistics archive to show *productivity data*[1] *(SFNI, 2011)*

[1] *Swedish: Bonitet.* Which is to evaluate the earning capacity of any natural resource

per region (among many other types of statistics). Another institution, The Swedish Forest Soil Inventory, also carries out long-term monitoring of the permanent sample plots of the SNFI. The two programmes are joined in cooperation within the Swedish National Inventory of Forests (Swedish Forest Soil Inventory, 2014).

| **Field Layer Types** |
|---|
| Tall herb type |
| Low herb type |
| Soil without field layer |
| Broad-leaved grass type |
| Thin-leaved grass type |
| Sedge-horsetail type |
| Bilberry type |
| Cowberry type |
| Crowberry-heather type |
| Poor dwarf-shrub type |

Table 5.1: Field Layer Types (Hagglund and Lundmark, 2004)

SNFI's productivity data is based upon the development of a method, or system for site quality classification based on site properties and are used to indicate *site productivity* (Hagglund and Lundmark, 1987, 2003 and 2004) . The method takes information on climate, soil mineral type, soil type and soil moisture and type of vegetation as primary inputs for its productivity indication output. The type of vegetation is important because it reflects soil fertility. Before establishing an output a field layer type classification (see table 5.1) has to be made. For this purpose the method involves a flowchart that assists the rater to determine the field layer type class for the area currently being monitored. One type of output could be presented as a map of the dominant field layer class for a particular area in Sweden, as shown in figure 5.1 (Dept. of Soil and Environment at SLU, 2012). (Reprinted with the permission of professor Lars Lundin at SLU, 2014-12-03).

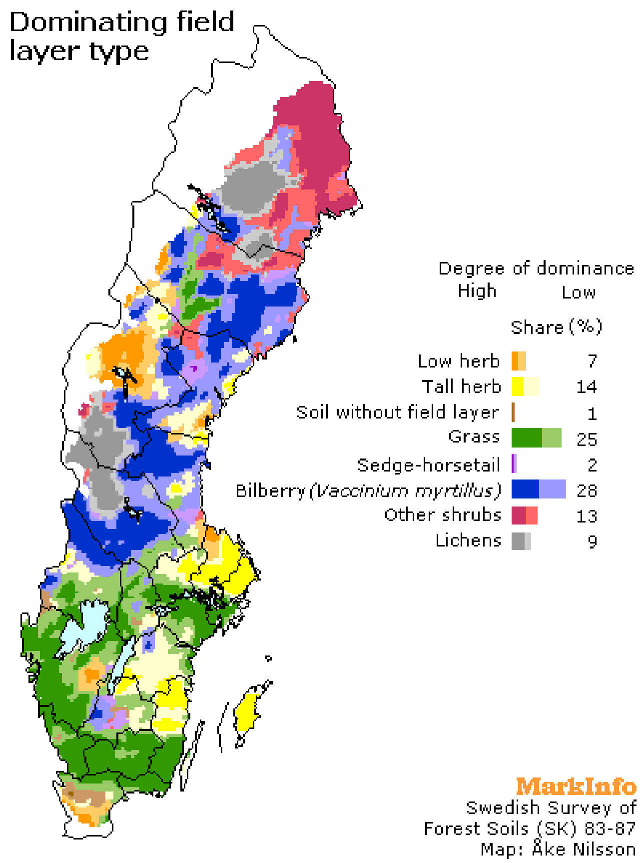



Figure 5.1: Dominant field layer class for a certain area (Reprinted with permission)

It is proposed that these field layer type classifications be used as one input when monitoring vegetation. The motivation for this is that, firstly, it would enable generalisations to be made between the same classes of field layer types. Secondly, the data will frequently be updated; and thirdly, data can be publicly accessible. The common sampling unit in the first three studies was a sub-plot frequency point quadrat grid consisting of 100 squares.

The underlying hypotheses for conducting the $1^{st}$ to the $3^{rd}$ field experiment can be summarised as follows:

1. Investigate if there is a correlation between the dry weight of the root and the dry weight of the shoot (i.e., the aboveground part of the

plant) in the *very special environment that exists on railway embankments*. The fulfilment of the environmental requirements (see section 3.2) is important for the root to shoot ratio. The environment on a newly laid railway embankment is initially very harsh for plants; however, this will gradually degenerate. As the ballast is ground down to fine stone powder and becomes polluted with organic litter from plants, the environment will, over time, become easier to invade, Sargent (1984). Nonetheless, the average railway embankment offers a relatively poor environment for most plants. In order to investigate if special root to shoot ratios do exist on railway embankments, a sampling process needs to be carried out by harvesting plants and excavating roots by hand. It has been shown in several publications that there are correlations between roots and shoots (see section 3.4).

2. If a correlation does exist, then root weights can be estimated by sampling shoots as part of this study. With regard to harvesting, it is the excavating of roots that is especially time consuming. Another major consumer of time is the activity of rinsing and cleaning the excavated roots in the lab. The motivation for these steps is that a camera sensor can only monitor the shoot. However, the shoot is only a part of the total biomass, so in order to estimate the biomass growing on defined areas of a railway line, the relationship (ratio) between root and shoot is vital. The process also investigates how vegetation is monitored and measured by humans. In order to transfer such a process, either fully or partially, into a system, the investigation and practical field experiment need to serve as an initial knowledge base.

3. The next step is to investigate if there is a relationship between aerial plant cover and the shoot part of the plant. This is supported by, for example, Rottgermann et al. (2000) and Krebs et al. (2003). The latter investigated four types of graminoids in a soil type with low nutrients and low moisture, and at a site that represented a transition stage from an open pioneer plant community to a sandy dry grassland. Jonasson (1988) showed that the number of point intercepts correlates with biomass. It is likely that new- to average-age railway embankment environments have similar conditions, i.e., low nutrient and low moisture. If the first hypothesis holds true, then there is a re-

lationship between the plant cover and the full biomass of the plant. In order to fulfil this, as well as document the work mentioned above, images need to be acquired.

The attribute cover (see section 4.3) was selected because of the similarity of on-site cover data judged by humans and cover data represented in an image and judged, or interpreted by machines. The on-site cover and cover seen on an image should be mapped. In field experiments 1 to 3, the camera was set up using a tripod (see figure 5.2), at a height of 1.6 m vertically above the ground to capture a nadir view of the trackbed.

Figure 5.2: Camera's vertical setup

## 5.3 1$^{st}$ Field Experimental Site: Falun - Grycksbo, Sweden

The first field experiment (FE1) was conducted over the course of two days on the 27 to 28 June 2011 at two different sections along the railway line between Falun and Grycksbo, WGS 84 decimal (lat, lon) coordinates 60.6657, 15.5437 and 60.6671, 15.5418, respectively (see overview of the railway section in figure 5.3). The general character of the experimental location is described in appendix E. It describes the field layer type and tree type at this site.

Weather conditions during both days: total cloud cover[2]: 1, sunny, almost clear. 25 to 27°C. Dry conditions.

Figure 5.3: Overview Grycksbo: Railway section under investigation

### 5.3.1 Motivation for Conducting the 1st Field Experiment

When transferring vegetation (plant) monitoring activities from humans to machines, it is rational to start by investigating how monitoring is traditionally carried out. This has been done by biologists for the past several hundreds of years and several methods are available. Questions that arise are: what attributes are typically used when monitoring plants? Which recognised methods are available? Such questions are dealt with in the chapters Chapter 3 and Chapter 4.

With regard to plants, the railway embankment environment generally comprises various herbs, including different types of grass, and woody (ligneous) vegetation. In order to identify and/or classify these plants, a representation system was set up.

Overall, two different main vegetation classes (strata) were chosen. These classes were chosen based on the previous experiences of administrators from STA. Huisman (2001) reported that Deutsche Bahn in Germany also suggested different classifications for railway embankments and their surroundings in Germany (see table 5.2).

[2]The appearance of the sky when 1/8ths of the sky is covered with clouds. SYNOP (Surface synoptic observations) numerical codes called FM-12 by World Meteorological Organization (WMO) is used here

| Grade of Coverage | Side Areas | Track Area |
|---|---|---|
| Zero | *Few or low-growing species* | *None or occasional plants* |
| Low | *Incipient vegetative cover, < 30%* | *< 2% coverage* |
| Medium | *< 50% coverage* | *< 10% coverage* |
| High | *> 50% coverage* | *< 20% coverage* |

Table 5.2: Suggested vegetation inventory classes by Below, Deutsche BahnHuisman (2001)

During a workshop held in 2011[3],it was found that the classes (mentioned in table 5.2) are still in use by Deutsche Bahn. The current class of a particular inspected area is estimated subjectively by the inspector of that area of railway.

This approach, which quantifies coverage into different classes is appealing. However, it suffers from the same problem of subjectiveness, as described in section 1.2, and will most likely result in different judgements depending on which person is currently carrying the inspection. In addition, there is no distinction between different species of vegetation. While different species have different abilities or properties (e.g. types of reproduction, growth speed, whether they are annuals or perennials, and type and size of root system), it would be beneficial to some extent to be able to segment the plants into classes according to species, genus or higher classification (see table D.1).

### 5.3.2 Method

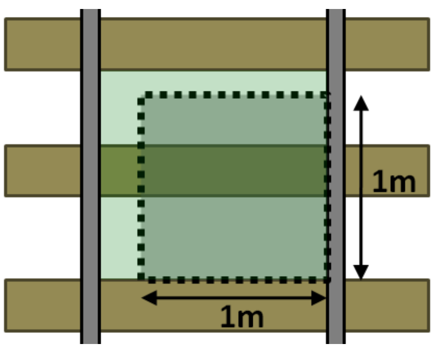
Two different segments of railway were chosen subjectively while walking along the track. A segment is defined as an arbitrary section of railway in which sample plot positions are then randomised. The segments were classified as Herbs-Sparse and Herbs-Dense, respectively. Within each segment five sample plots were randomised along the rails making a total

3

**Workshop** *Sustainable Management in Rail Environment* held by the Swedish University of Agricultural Sciences, and supported by the Swedish Transport Administration, on the 27-29 April 2011 in Alnarp, Sweden.

of ten sample plots. Within a *sample plot* is a *sampling grid frame* that measures 1m$^2$ (a 1-by-1 m area), as seen on the right-hand side of figure 5.4. The *sampling area* is defined as the light-green area, also seen in figure 5.4 (left-hand side). The position of each sampling frame in each sampling area was selected at random[4], once.

The sample grid frame was centred in between the rails and the upper part of the lower sleepers, as seen in figure 5.4a. The sampling frame was divided into 100 squares, each defining 1% of the total sampling area (see figure 5.4b).

a) 

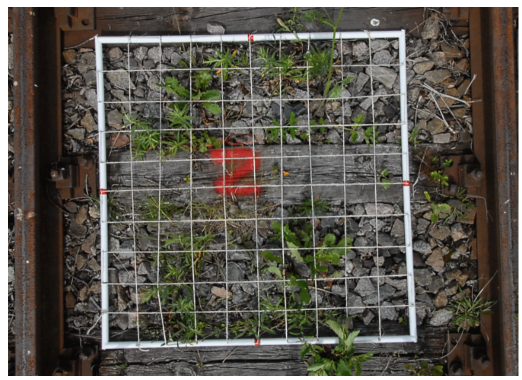

b) 

Figure 5.4: a) Sample plot area, and b) point quadrat grid frame containing 100 sub-plots

A representative railway segment (stratum) was selected to represent sparse herbs and dense herbs, respectively. Five sample areas (each 1 x 1 m) were randomly selected within each area using a simple random selection method. In total the two strata together contained 10 sampling areas. Each sampling area was marked with a unique and identifiable identification number on the centre sleeper using permanent ink.

First, a subjective judgement was performed by visually estimating plant cover in the corresponding square. Note that no advance instructions of how to judge were given. Instead, each and every person had to, on first sight, judge the total coverage as a percentage of the area occupied by

[4]Eight frame positions were in the universe: Upper left, upper middle , upper right, centre right, and so on clockwise.

the 1 x 1 m frame. One should also note that the sampling frame was not placed over the vegetation at this time.

Two methods of judging cover were applied: *Aerial foliage cover* and *sub-plot frequency* Elzinga et al. (1998), Elzinga et al. (2001) and Coulloudon et al. (1999) (see chapter Chapter 4). During this study the following protocols were applied:

*Aerial Foliage Cover (AFC):* Only the vegetative parts of the plants were measured, i.e., no spacings (see left-hand side of figure 4.4). If a part of a plant appeared in a square in the grid, then its total cover area was estimated and kept in mind. Each estimate could not exceed 1%, because a grid square equals 1%. For example in the sub-area depicted in figure 5.5 this method would estimate the cover in each grid square as seen in table 5.3a. The individual grid square estimates are then summed up. Out of a total of 9 grid squares (i.e., 9%), the cover would be 2% (out of a possible 9%).

a)

| | | |
|---|---|---|
| 0 | 0.70 | 0.50 |
| 0 | 0.35 | 0.20 |
| 0.25 | 0 | 0 |

b)

| | | |
|---|---|---|
| 0 | 1 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 0 |

Table 5.3: Estimating cover (%) using a) AFC, and b) sub-plot frequency

*Sub-plot Frequency:* The method used here is a consistent way of estimating cover. It was carried out using the strategy adopted from the sub-plot frequency method. If part of a plant was in a grid square, then it counted as 1%; if not, then it counted as 0%. Only vegetative parts of the plants were measured, i.e., no spacings (see figure 4.4b). In the same example as above (figure 5.5), a person who makes use of this method would certainly estimate the cover in each grid square as seen in table 5.3b. Out of a total of 9 grid squares, the cover would be 5% (out of a possible 9%). The method will most likely overestimate the true cover, especially if the grid is coarse.

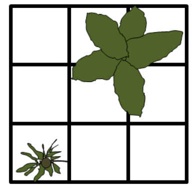


Figure 5.5: Subset of a sub-plot grid, or point quadrat grid

The heights of all plants belonging to each square were measured using a folding rule. Although every plant were measured, these measurements were not mapped for each plant individually, i.e., it is not possible to find out which recorded height belonged to which plant from the acquired images. Mapping every recorded height to individual plants (e.g. by drawing sketches and plotting every plant during the field experiment) were ruled out because of time constraints, that included the limited availability of a security and safety officer.

After measuring the heights, a grid frame was placed over each of the ten sampling areas, as seen in figure 5.4, and the cover was estimated using the two different protocols, as discussed above.

Thereafter, when the cover was estimated by the raters, two images were acquired vertically above each plot. The centre of each image was approximately the centre of the sleeper in the middle of sampling area (figure 5.4).

All plants belonging to each of the first five (sample plot numbers 1 to 5) 1x1 m areas were then harvested and the roots excavated. Plants were put into plastic bags according to their species. All plants from one sample plot were then put into a bigger plastic sack marked with the sample plot number. The same procedure was applied for sample plots 6 to 10; however in these five plots no roots were excavated primarily because of time constraints.

**Drying of plants and weighing procedure**

All ten plastic bags were then transported to the lab and put into a 20°C storage room during the night. Next day, each plant was identified. The soil was then washed away from the roots, and the roots were separated from the shoots (the cuts were made at the soil line). If there were several plants of the same species they were put into the same paper bag. Thus, in addition to the species name and a serial number, the number of individual plants was noted on the paper bag. Note that from this step, each plant collected from sample plots 1 to 5 was separated into root and shoot, and placed in separate paper bags. Typically, each paper bag was marked as: <species name>, <root, or shoot>, <location>, <sample plot number>, <number of individuals>, <unique serial number>. The date, paper bag type, and so on were then coupled with the serial number.

All paper bags and their contents were left to dry in ovens set to 105°C overnight. Next day, the plants were left to cool off in a Exsickator (i.e., a moisturefree environment), so that neither the plants, nor the paper bags would take up water from a humid environment. After cooling off for approximately 15 minutes, the open paper bags and their contents were weighed on a *Mettler PM460* scale with a precision down to milligrams. The gross weights were recorded.

The paper bags were of three kinds, and, as each individual bag may have slightly different weights, a sample of 10 to 20 bags of each kind was dried in an oven and weighed. The arithmetic mean weight for each kind of paper bag was then subtracted from the gross weight to get the net weight of the contents. All the dried plants were then stored in case of later usage.

### 5.3.3 Results and Conclusions of the 1st Field Experiment

A detailed analysis from the drying of plants is reported in 9.

Total data acquired in this field experiment from the ten sample plots: Vertically, from above the ten sample plots, 20 VIS images were acquired. After that, a total of 353 plants were harvested from the ten sample plots: 175 plants were excavated with both roots and shoots from sample plots 1 to 5; and 188 plants were harvested from sample plots 6 to 10, but this time only the shoots were excavated. See detailed weight data in Appendix A.

The collected data from the (human) visual estimates are presented in tables 5.4 and 5.5.

| Visual Cover Examination | SP1 | SP2 | SP3 | SP4 | SP5 | SP6 | SP7 | SP8 | SP9 | SP10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Rater #1 | 35 | 40 | 30 | 15 | 15 | 20 | 15 | 20 | 35 | 25 |
| Rater #2 | 25 | 30 | 20 | 5 | 10 | 10 | 5 | 10 | 30 | 20 |
| Rater #3 | 10 | 25 | 20 | 5 | 10 | 15 | 10 | 10 | 35 | 10 |

Table 5.4: Non-guided visual cover estimation

| Visual Cover Examination | SP1 | SP2 | SP3 | SP4 | SP5 | SP6 | SP7 | SP8 | SP9 | SP10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Aerial Foliage Cover | 15 | 16 | 15 | 8 | 10 | 20 | 17 | 10 | 36 | 14 |
| Sub-plot frequency | 19 | 23 | 25 | 14 | 17 | 33 | 24 | 21 | 77 | 53 |

Table 5.5: VE cover estimates using a sub-plot grid quadrat

The images used in the analysis in section 6.6 were laypersons' estimates of the plant cover in each image. Another analysis of visual estimates was made by domain experts is presented in section 6.7. The heights of all plants belonging to each sampling square were measured. These measurements were not mapped for each plant individually.

The visual estimation of plant cover produced results that gave strong indications of a large variation. Thus, it most likely can be considered an imprecise method. Although the number of persons judging were small and the field experiment was limited, it does verify the disadvantages stated in section 1.2.3; namely, that the results of a visual inspection are highly subjective, depending on the experience and knowledge level of the person who is carrying out the inspection.

In some cases, it is very hard to identify individual plants, especially when they grow tight together in tufts, clusters, or clumps. It is important to stress that it is not the purpose of this thesis to identify individual plants rather than tufts, clusters, or clumps. As a result, the number measurements of heights might not reflect the actual number of individuals. A more practical way, when individuals cannot be identified, could be to count plant clusters instead of individual plants.

During the first visual estimation it became apparent that two different methods were used to estimate cover. These methods were aerial foliage cover, and sub-plot frequency (see sections 4.3.1 and 5.3.2). This will, of course, have an impact on the results, because the same area would be quantified differently. Consequently, two protocols were created, so that it would be clear how to estimate cover, depending on which of the two methods were used.

The benefit of the protocol used for estimating according to the *aerial foliage cover* method is that it will be quite near the ”true” cover value. The drawback is firstly that the person has to keep float numbers (i.e. fractions of plant cover) in the back of his/her head in the continuous summation from plot to plot. This makes the method uncertain over time. Secondly, it takes a trained person to make as correct cover estimations as possible.

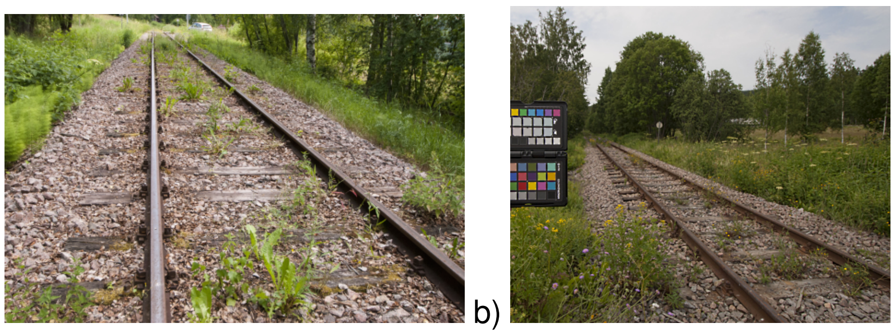
The benefit of the protocol used for the sub-plot frequency method is that it is easy to use. If part of a plant is found within a grid square, the total is increased by 1%; if not, then 0% is added. The summary will be more like an integer enumeration. The drawback of using this method is that it will almost always deliver an over estimation when compared with the ”true” cover value. This over estimation will decrease with more and smaller grid squares.

When comparing these two protocols, the latter (the protocol used for the sub-plot frequency method) is preferable because it offers consistency. The aerial foliage protocol is estimated (not measured) from fractions of a sub-plot. On the other hand, the sub-plot protocol, even though it produces over estimations compared with the ”true” value, is still strictly speaking a form of measuring, thanks to the rigid protocol of counting (see 4.3.2).

## 5.4 2$^{nd}$ Field Experimental Site: Falun - Grycksbo, Sweden

This field experiment study was conducted over the course of two days on the 6 to 7 July 2011 in one area (near Bergsgården’s old railway station) along the railway line between Falun and Grycksbo , from the starting point at WGS 84 decimal (lat, lon): 60.66841, 15.53918 and 65 m in SE direction along the railway tracks (see figures 5.6a and 5.6b). The general field layer type of the experimental location are described in appendix E.

Weather conditions during both days: total cloud cover: 3, sunny, almost clear. 26 to 29°C. Dry conditions.



a) b)

Figure 5.6: Overview of the railway section

The purpose of this particular study was to obtain data from the Bilberry and Grass field layer class, see section 5.2. Some woody plants (deciduous trees) were also present, whilst grass was almost absent.

This railway section (mainly containing herbs) was selected subjectively by a domain expert. In advance, it was decided that every second sleeper should constitute the centre of a sampling area (see the light-grey area in figure 5.7). A starting point on the tracks was selected randomly, and, from that starting point, 50 sampling areas were investigated along this railway section.

Each sample area was defined as the ballast area on each side of a central sleeper, inclusive of the sleeper, as shown by the light-grey area in figure 5.7. Each sample frame was defined as a 1 by 1 m quadratic area. The position of the sampling frame in the sampling area was selected randomly, once. Eight frame positions were in the universe: Upper left, upper middle , upper right, centre right, and so on clockwise.

The randomisation of the quadrat position ended up in the position lower right, which is at the upper boundary of the right rail and the lower sleeper', as shown by the dark-grey area in figure 5.7. This sampling frame position was then applied for all 50 sampling areas.

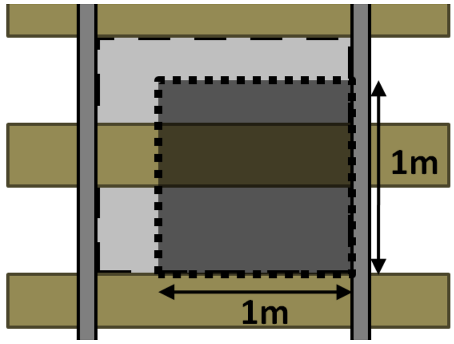



Figure 5.7: Sample plot area at 2nd field experimental site.

The remaining study was conducted in the same manner as the first field experiment (see section 5.3, except that no plants were harvested, nor were any shoots excavated. Additionally the distance between sleepers as well as sleeper widths were measured using a laser distance meter. These measurements were also transferred to support the fourth field experiment, see section 5.7.

### 5.4.1 Results and Conclusions of the 2nd Field Experiment

In this study of 50 sample plots, 100 VIS images were acquired, capturing a nadir view over each sample plot. Visual estimates from the two domain experts and the measuring of plant heights were recorded. No plants were harvested in this study.

The purpose of this field experiment was to acquire data from the field layer class *Bilberry and Grass* (see section 5.2). For this purpose, a railway section containing herbs was selected. From a randomised starting point on the rails, 50 sampling areas were investigated along this railway section.

## 5.5 $3^{rd}$ Field Experimental Site: Älvdalsbanan, Oxberg

The purpose of this study was to obtain a large amount of data from one particular vegetation class: conifer woody plants. The dominant field layer class in Oxberg can be characterised as *Lichens* (see section 5.2). More information about field layer type and tree types at this experimental site can be found in in appendix E.

Figure 5.8: Overview of the sampled railway section (Oxberg, Sweden)

The field experiment was conducted over the course of four days on the 11 to 12 August 2011 (monitoring and imagery) and from the 18 to 19 August 2011 (excavating root and shoots) of sampled railway a section of railway between Mora and Märbäck, from a starting point in Oxberg at WGS 84 decimal (lat, lon): 61.13397, 14.17138 and continuing about 60 m in N direction along the rail (see railway section overview in figure 5.8).

Weather conditions during 11th to 12th August: total cloud cover: 5. 22 degrees Celsius. Dry conditions. Weather conditions during 18th to 19th August: total cloud cover: 7. The temperature was 15ºC. Occasionally showers: 80 (light rain showers).

The study was conducted in the same way as the former field experiments. Images documenting sample plots numbers 200 to 204 were acquired (see figure 5.9). Both roots and shoots were harvested from all the plants. The only difference from previous field experiments was that when the plants were put into bags, they were sorted not just in terms of species, but also in terms of size. Consequently, every bag contained plants of the same species and about the same size.


Figure 5.9: Sample plot no. 200 a) before, and b) after excavation

The environment was in the Lichens field layer class; thus, it was different from the field experiments in Grycksbo, Sweden (see sections 5.3 and 5.4), in the sense that it was a relatively harsh environment. Mostly graminoids and conifer woody plants were growing on the railway embankment (see figure 5.10).


Figure 5.10: Conifer woody plants, in plot 202 (Oxberg, Sweden)

Figure 5.11: Documentation of excavated plants in lab (Picea abies)

### 5.5.1 Results and Conclusions of the 3rd Field Experiment

In this study of five sample plots (ID nos. 200 to 204), 20 VIS images were captured from a nadir view.

In total,197 plants (mostly conifer woody plants and graminoids) were harvested from the five sample plots (ID nos. 200 to 204).

The harvesting activity included excavation of both roots and shoots. For the analysis of the collected data see section6.10.

The purpose of this study was to obtain data from the woody plants vegetation class. It should be noted that none of the plants were higher than the distance from the ground to a train's axles, i.e., the plants are constantly being cut down. It should be noted that it took one day for one man to excavate the five sample plots and put the plants into bags. Some of the bigger plant individuals, such as Picea abies, took 45 minutes (each) to excavate because of their widespread root system.

## 5.6 4$^{th}$ Field Experimental Site: Vetlanda

Images acquired were collected during daytime on two separate occasions (in June and August 2013) along a 700 m long section of railway located outside Vetlanda, Sweden, WGS 84 decimal (lat, lon) coordinates 57.42992, 15.17194 and 57.428633, 15.183137, respectively (see railway section overview in figure 5.12). Information about the field layer type and tree type at this experimental site are described in appendix E.

Figure 5.12: First part of the Vetlanda railway section

A DSLR Nikon D90 camera mounted on a stand, which in turn was mounted on a trolley, was used to acquire the relevant images in visible spectrum (400 to 700 nm). No artificial lighting was used. All the images were of high resolution (4288*2448 pixels) and were saved as RAW files in the RGB colour space. This is because, images stored in RAW format can be post processed, if necessary. The camera was mounted at a height of 2.25 m above the ground, pointing slightly downwards and forwards in the driving direction, 65 degrees below the horizontal line (see figure5.13). Weather conditions at both data collection sessions were sunny, with few cumulus and about 25ºC in the shade. The ground was dry.

Figure 5.13: Trolley with mounted camera

A Tamron Ultra Wide Angle Lens SP AF10 - 24mm F/3.5-4.5 Di II LD Aspherical [IF] was used and the focal length was set to 10 mm, resulting in a wide angle of 100°in width. Every $10^{th}$ sleeper was marked with a number, 0 to 89, with yellow spray paint. Each sample area (or sample plot) contained five rectangular ballast-filled areas in between six sleepers (see figures 5.14a and 5.14b). At each sample area (i.e., at every fifth sleeper) the camera was wirelessly fired using an Aion Wireless Timer Remote Controller. The entire 700 m railway section was documented in this way.

a) 

b) 


Figure 5.14: Vetlanda sample plots a) ID 4, and b) ID 22

As already stated, data acquisition was carried out twice (along the same section) to be able to capture growth of vegetation during the summer months in Sweden. In all, 354 images were acquired from both sessions. Out of 179, only 176 images from June's acquisition and 178 images from August were deemed usable for further analysis. The remaining images in both the cases were unusable and were therefore excluded.

## 5.7 $5^{th}$ Field Experiment

**Motivation**: The general sampling area, as mentioned before, was set as the two rectangular ballast areas on either side of a central sleeper, inclusive of the sleeper, shown as the light-grey area in figure 5.7. It is likely though that not every sleeper will be found when analysing the image by hand. This might be because ballast and/or vegetation cover the sleeper to such an extent that the sleeper's edge cannot be found. In such cases, the position of the sleeper can be estimated by using the standard distance (i.e., the spacing) in between sleepers. For that reason (when data is incomplete due to loss of information, or omitted because of various sampling procedures) data has been collected to be able to generalise the distance between sleepers as well as sleepers widths.

The regulations of the Swedish Transport Administration Banverket (1995) state that the spacing in between sleepers should normally not exceed 650 mm on straight tracks, as well as in curves of radius, $R$, such that $R \geq 500m$. The spacing should not exceed 600 mm in curves of radius, $R < 500m$[5]. This applies to newly laid sleepers so when it comes to sleepers mounted possibly 50 years ago there were no standardised maximum, or minimum spacings to be found in the literature.

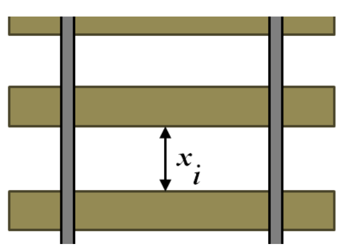



Figure 5.15: Sleeper spacing

[5] Curves having a radius, $R < 150m$ are normally not used Banverket (1996).

Due to the lack of standardisation As a result of this lack of standard distances, a collection of 410 sleeper spacing measurements was sampled as follows. Firstly, the sampling areas of the wooden sleepers were chosen according to availability and security. Secondly, when visiting each area a systematic sample was acquired i.e. at a chosen point on the railway an integer was randomised to describe how many sleepers lay ahead of the chosen starting point of the sampling. The sampling interval was randomised as once in every fifth sleeper. Consequently, from the randomised starting point, every fifth spacing was measured using a laser distometer Leica Disto D8. The spacing was measured from edge to edge (see figure 5.15). Also during the second field experiment 150 wooden sleeper widths were collected.

### 5.7.1 Results and Conclusions of the 4$^{th}$ Field Experiment

*Sleeper spacing distances:* All sampled data from this investigation are presented in the histogram (see figure 5.16). The mean is represented as a solid line and the median as a dashed line. The sample mean $\overline{x}$ of the $n$ sleeper spacings was calculated to be 0.486 m, with sample standard deviation $s = 0.117$ m. At 95% confidence level for the sample mean results in a confidence interval of 0.486$\pm$ 0.011, or [0.497; 0.474]

*Sleeper widths:* Concerning the sleeper widths, at 95% confidence level for the sample mean $\overline{x}$=0.231 metres results in a confidence interval of 0.231$\pm$0.579, or [ 0.225; 0.237]

a)

| Sleepers spacings | | |
|---|---|---|
| Number of samples: | 410 | |
| Median: | 0.500 | m |
| Mean: | 0.486 | m |
| Std.Deviation: | 0.117 | m |

b)

| Sleeper widths | | |
|---|---|---|
| Number of samples: | 153 | |
| Median: | 0.23 | m |
| Mean: | 0.231 | m |
| Std.Deviation: | 0.037 | m |

Table 5.6: a) Spacings between sleepers. b) Sleepers widths

Systematic sampling was conducted along railway tracks in the neighbourhoods of Borlange, Grycksbo and Oxberg, where 410 sleeper spacings

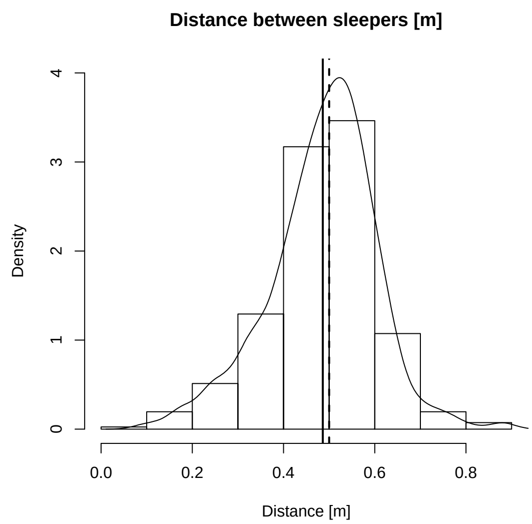


Figure 5.16: Data - Sleeper spacings

were measured. 153 sleeper widths were measured in total (see table 5.6). These measurements are to be used when a sleeper cannot be found in an image when trying to find the sampling area boundaries, as seen in the light-grey area in figure 5.7. Results of this study indicate that a sampling area can be computed using the estimates for the mean of the sleeper widths and sleeper spacings.

## 5.8 Indoor Laboratory Experiments

### 5.8.1 Motivation for Laboratory Experiments

It was assumed that woody plants are regarded as being more undesirable when growing on the embankment than herbaceous plants and graminoids (i.e., grass) (see section 3.5). This is because, after establishment woody plants produce a large biomass and are more difficult to control. This assumption was confirmed by Lindstrom (7 Sept. 2011, Personal interview) and Lundh (28 April 2011, Personal interview), who also stated that ”only graminoids constitute a threat to the woody plants, but that is only during the tree seedling establishment phase, i.e. during the first year”. After this phase, the graminoids can no longer out-compete woody plants in the competition for the uptake of water and essential nutrients.

As a consequence, it is desirable to detect, identify and characterise Norway spruce, Scots pine and birch trees (including Downy birch and Silver birch) (see section 3.5). The hypothesis in this case is that they can be detected and characterised using machine vision. Here, the system is presented in advance with images of each woody plant and has an opportunity to identify and learn their signatures, i.e., making use of supervised machine learning. For this reason, additional data has to be acquired. The purpose of data acquisition is to gather a set of images that represent each woody plant type, which will then form the base for a supervised learning process (see the analysis and results from the acquired data in chapter 10).

The rationale for starting with laboratory experiments is as follows: if a species cannot be classified by leaf recognition in a controlled laboratory environment, then it is unlikely that it can be classified outdoors in an uncontrolled environment.

#### Data Collection of Leaves

One approach to reducing the complex problem of identifying woody plants on railway embankments is to identify the plant species by their leaves. Possible features which could be used to identify the species, genus, or family are shape, texture, and colouration.

The data collection was based on collecting leaves from the most frequently found deciduous woody plants and to obtain conifer seedlings from the most frequently found conifer plants. Official forest statistics in the table *Standing volume for different tree species by diameter class - All land-use classes 2006 -2010* from the Swedish National Forest Inventory reported the most frequently found species in Sweden. Those having a diameter (at chest level) of 0 to 9 cm are listed in table 5.7.

| Familia | Genus | Species | English | Swedish | Species Composition (%) |
|---|---|---|---|---|---|
| Pinaceae | Pinus | sylvestris | Scots Pine | Tall | 21.8 |
| Pinaceae | Picea | abies | Norway Spruce | Gran | 35.0 |
| Pinaceae | Pinus | contorta | Lodgepole Pine | Contortatall | 2.0 |
| Betulaceae | Betula | pubescens | Downy Birch | Glasbjörk | 33.4 |
| Betulaceae | Betula | pendula | Silver Birch | Vårtbjörk | |
| Betulaceae | Alnus | incana | Grey Alder | Gråal | 2.8 |
| Betulaceae | Alnus | glutinosa | Alder | Klibbal | |
| Salicaceae | Populus | tremula | Aspen | Asp | 1.0 |
| Rosaceae | Sorbus | aucuparia | Rowan | Rönn | 1.3 |
| Fagaceae | Quercus | robur | Pedunculate Oak | Ek | 0.6 |
| Fagaceae | Fagus | sylvatica | Beech | Bok | 0.2 |
| Salicaceae | Salix | caprea | Goat Willow | Sälg | 1.0 |
| Rosaceae | Prunus | padus | Bird Cherry | Hägg | < 1.0 |
| Aceraceae | Acer | platanoides | Norway Maple | Lönn | < 1.0 |

Table 5.7: Tree frequency in Sweden

The data in table 5.7 together with the suggested woody plants in section 3.5 helped in the choice of which data to collect. Time constraints meant that some of the species were excluded from the data collection, e.g. the Fagaceae Fagus sylvatica (beech).

### Experimental Setup

All images were acquired indoors in a photo laboratory. Images were acquired using a Nikon DSLR D90 camera (3872x2592; 380-780 nm).

The DSLR was mounted on a tripod stand (to maintain uniform distance) placed 160 cm vertically above the object table to capture a nadir view (see figure 5.17).

Figure 5.17: The indoor lab setup

The leaves and conifer plants were illuminated as shown in table 5.8.

| | Leaves session | | Conifer plants session | |
|---|---|---|---|---|
| | Colour Temp. (Kelvin) | Illuminance (Lux) | Colour Temp. (Kelvin) | Illuminance (Lux) |
| Top | 2200 | 455 | 2200 | 512 |
| Left | 2100 | 700 | 2100 | 777 |
| Right | 2300 | 600 | 2200 | 600 |
| Under | 1900 | 185 | 1900 | 58 |

Table 5.8: Measured average incoming light at the object

a) b) c) d) e)

Figure 5.18: Examples of conifer plants and leafs on the object table

In order to be able to correct the white balance, or grey balance, a grey card (QPcard 101 of size 142 x 40 mm[6]) was placed in each acquired image. The QPcard contained three fade-resistant fields in white (CIE Lab 95*0*0), mid gray (CIE Lab 48*0*0) and dark gray (CIE LAB 35*0*0). In addition, each card contained a millimetre scale of about 30 mm (as a reference). The grey colour field on the grey card reflected 18% of the light making it useful for exposure purposes.

## 5.9 Conclusion and Discussion

In all the outdoor studies (except for the sleepers study) a sampling frame constituting a square of 1 x 1 m was used. The frame was divided into 100 sub-plots (each sub-plot area was 10 x 10 cm), and thus each such square represented 1% of the frame area. This sampling frame was chosen in preference to the popular Daubenmire frame, which constitutes a rectangle of 50 x 20 cm, equalling 100 cm$^2$ Daubenmire (1959). This decision was based on the belief that it would be easier and less time consuming to count occurrences.

The question arises: "Which is the best form of sampling frame?" This may depend on how the sampled vegetation is distributed spatially. Most often, plants are aggregated in clumped distributions (see section 4.2). For this kind of distribution, it has been shown that a rectangular sampling frame reduces the number of zero counts as well as counts that are very high Elzinga et al. (1998). The same source also recommended that the longer side of the rectangle frame have a measure bigger than the mean distance between clumps.

It has been observed during these studies that the vegetation on the track bed/ embankment is likely to be patchy in character, i.e., with clumped (aggregated) populations.

In cases of clumped distribution it has been proposed that transects, not quadrats, should be treated as the sampling unit, because "... *the transects will intersect several clumps of the population, this ensures much of the variation will be incorporated within each sampling unit. If individual*

[6] http://www.qpcard.com/ (Retrieved: 2012-08-21)

*quadrats are treated as the sampling units, most of the variation will be between sampling units."Elzinga et al. (1998, p.111)*

The indoor experiments resulted in approx. 700 images were acquired compromising Pinus sylvestris, Picea abies, Betula pubescens, Betula pendula, Alnus incana, Alnus glutinosa, Populus tremula, Sorbus aucuparia, Quercus robur, Salix caprea, Prunus padus, and Acer platanoides. The analysis of this data is presented in chapter 10.

# Part IV

# Manual Assessments of Terrestrial Vegetation

This part includes investigations of how to manually assess the extent of terrestrial vegetation. In the context of this work, manual assessments of terrestrial vegetation refers to the methods used by humans to measure, or estimate the extent, or amount of plants at a specified spatial location. This could, for example, be one or several areas, or be along one or several line transects (i.e., pre-determined vectors).

Vegetation assessments within railway maintenance are largely carried out manually by visually inspecting the track on-site, or by looking at video clips collected by maintenance trains, or trailers as they run along the track. Hence, it was deemed important to evaluate human assessment abilities in evaluating cover by visual estimates, and even further to see if different raters agree upon the estimates reported by each other. Several existing publications on the manual assessment of the extent of vegetation have come mainly from the ecology and botany domains. These papers have dealt with cover, frequencies, densities, and the biomass of certain species. No published investigations have been found in the literature review about assessing vegetation on railway embankments.

In the railway domain, no vegetation is deemed to be desirable when found on embankments. Therefore, total: cover, frequency, density and biomass are of particular interest.

# Chapter 6

# Visual Estimates, Reliability & Raters Agreement

## 6.1 Recording Human Estimates

The data collection here refers to the assessments made by two or more raters. The assessments in the context of this thesis most often refers to the estimates of vegetation extent, e.g. counting the number of plant individuals, or estimating the plant cover. The assessors in each investigation made their ratings/assessments on-site, or from images, see description of field measurements in chapter 5. A short summary and description of the investigations made:

1. Thirteen laypersons conducted visual estimates from 10 nadir images on Grycksbo sample plot nos. 1 to 5 and 6 to 10. The results are presented in section 6.6. For further details refer to (Nyberg et al., 2013b).

2. Three domain experts conducted visual estimates from 10 on-site observations in Grycksbo on sample plot nos.1 to 5 and 6 to 10.. The results are presented in section 6.7. For further details refer to (Nyberg et al., 2013b).

3. Three laypersons made visual estimations of cover from 35 images. The methods used for these estimates were ACC, and AFC. The

results are presented in section 6.8.

4. Five national maintenance engineers from STA North, STA South, STA East, STA West, and Borlange Municipality conducted visual estimates from 51 nadir images. The methods used for these estimates were ACC and AFC. The results were presented in section 6.9.

5. Two domain experts conducted visual estimates from five images from Oxberg, sample plot nos. 200 to 204. The methods used for these estimates were ACC, AFC and SF. The results are presented in section 6.10.

6. Three domain experts made visual estimates from 12 onsite observations in Vetlanda on two different occasions. The method used for these estimates was ACC. The results on density are presented in section 6.11. For further details refer to Nyberg et al. (2014).

## 6.2 Levels of Measurement

After acquiring data a choice had to be made, whether to use parametric methods or non-parametric methods for the upcoming statistical analysis.

Initially the level of measurement has to be decided, i.e. *nominal* scale, *ordinal* scale, *interval* scale, or *ratio* scale (see table 6.1). These levels are sometimes also referred to as scales of measurement, and were originally developed and published by (Stevens, 1946).

The lowest measurement is the nominal scale, and the highest level is the ratio scale. In many cases, the variable affiliation is obvious, but as the level of measurement can be reduced, lower levels might be appropriate. For example, measurements collected on an interval scale may be reduced to an ordinal scale, but not the other way around. This could be that case if the researcher wants to divide the measurements into classes of numerical intervals.

A variable in a nominal scale (also called categorical) can only be assigned to a defined category if it belongs there. If it does not belong there, it is called a dichotomous variable. There is nothing in between these two, and these kind of variables cannot be ordered. In the ordinal scale (or ranking scale), the variables can be ordered in relation to each others, but there is no equal distance in between the ranks; so, for example, if we have the ranks: $A = 4, B = 2, C = 1$ , so $A > B > C$. It is only possible to conclude that A is larger than B and C, and that B is larger than C. It is not possible to conclude that A is four times as large as C. The letters might represent different levels of attitudes, such as measurements on the Likert scale. The numerical rank representation just provides a way of sorting the ranks. Central tendency measurements, such as different kinds of means (averages), or the standard deviation, can be computed (based on the rank values); however, they have no practical meaning. Instead, the median and mode are used as a measure of central tendency in the ordinal scale. In the nominal scale, only the mode gives a meaningful measure of central tendency.

A variable on the interval scale has all the properties that an ordinal variable has. In addition, the intervals between the values are equally spaced. For example, the time difference in between 4 and 6pm is the same as the difference between 10 to 12am. By way of another example, we have A = 4; B = 2; C = 1 , so A > B > C as before, but this time (because of the equal space) one can conclude that the difference between A and B is 2, which is the same difference as in between B and C, and the difference between A and C is 4. In the case of the ordinal scale there would not be meaningful to calculate differences.

A variable on the ratio scale (i.e., the highest level) has all the properties of the three lower levels. In addition, it always has a clearly defined zero point where none of the property being measured exists. For example, we have A = 4; B = 2; C = 1 , so A > B > C as before, but this time (because of the equal space) one can conclude that A is two times larger than B, A is four times larger than C, and that B is twice as large as C. The parametric methods require data to be:

1. normally distributed, at least approximately, in order to be valid. For example, the parametric t and F tests also have these underlying assumptions (Siegel, 1957).

2. the observations must be independent
3. the population data are assumed to be homoscedastic, i.e., having the same variance; and
4. when conducting an analysis of variance (ANOVA), the means of these populations must be linear combinations of effects, because columns and/or rows-the effects must be additive, and
5. because of the comparisons of means, the measures have to be additive i.e.- numerical.

If normality can be assumed, then parametric methods can be used; otherwise, the data analysis should be made using non-parametric methods. Parametric methods are considered to be more powerful than non-parametric methods. However, if one assumes that the data set in question is normally distributed, but in fact it is not, then parametric methods can be misleading. The power of a test is defined as the probability that the test will reject the null hypothesis when in fact it is false and should be rejected. A statistical test is more powerful if it has small probability of rejecting the null hypothesis ($H_0$) when $H_0$ is true, but a large probability of rejecting $H_0$ when $H_0$ is false (Siegel, 1957). In general, every parametric method has its equivalent non-parametric method.

The non-parametric methods deals with qualitative data on a nominal (or categorical) scale and data on a ordinal (or ranking) scale respectively.

| **Scale** | **Defining relations** | **Example of statistics** | **Appropriate tests** |
|---|---|---|---|
| Nominal | (1) Equivalence | Mode<br>Frequency | Non-parametric tests |
| Ordinal | (1) Equivalence<br>(2) Order | Median<br>Percentile<br>Spearmans $r_s$<br>Kendalls $\tau$ | |
| Interval | (1) Equivalence<br>(2) Order<br>(3) Ratio of intervals | Mean<br>Standard deviation<br>Pearsons $r$ | Non-parametric, or<br>Parametric tests |
| Ratio | (1) Equivalence<br>(2) Order<br>(3) Ratio of intervals<br>(4) Ratio of values | Geometric mean<br>Coefficient of variation | |

Table 6.1: Levels of measurement, (Siegel, 1957)

In the six investigations (in section: 6.6, 6.7, 6.8, 6.9, 6.10, and 6.11), *box plots*, or box-and-whisker plots, were used to describe data by displaying the spread of all the data points in each data sample (Tukey, 1977), (McGill et al., 1978). By convention they outline five values: the extreme values (The biggest and the smallest values in the data set, except outliers) visualised as so called *whiskers*, the upper and lower hinges (displays the quartiles, i.e. the 25$^{th}$ and 75$^{th}$ percentiles), and the median (i.e. the 50$^{th}$ percentile). In addition, outliers (if any) are often shown above and below the whiskers, see figure 6.1.

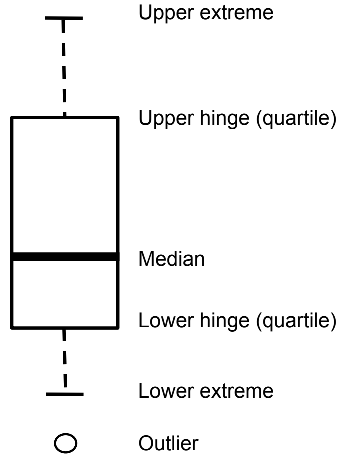
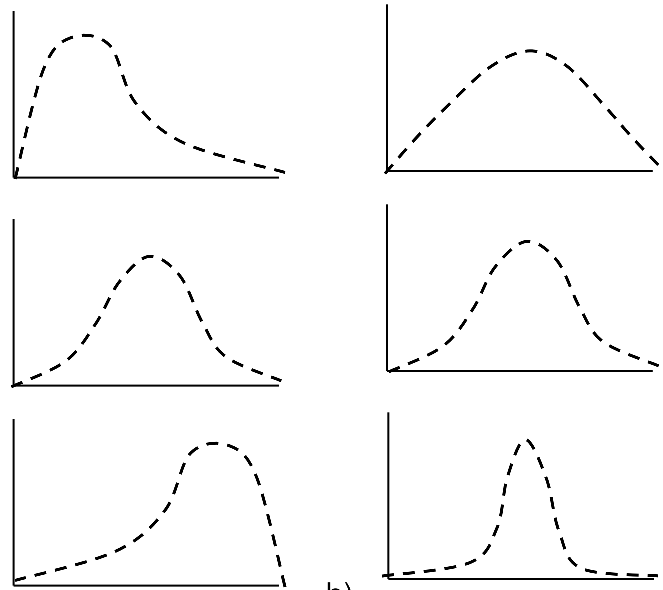



Figure 6.1: Boxplot

### 6.2.1 Data Transformation

When data are to be analysed the choice is to use either parametric methods (if the data set is assumed to come from a normal population), or non-parametric methods (if one cannot assume that the data is normally distributed, or if the number of samples is few) (See section 6.2). In some cases, data can be transformed from being non-normal to normal. This is a very common pre-analysis procedure, the purpose of which is to make it possible to analyse data. In such a case, the data is transformed by performing a mathematical operation on each data point (observation). These transformed observations are then used in statistical tests. In general a distribution can deviate from normal in two ways. Firstly, skewness, also referred to as the third standardised moment. Secondly, kurtosis also referred to as the fourth standardised moment (see figure 6.2). These are measurements of shape and are often used to test for normality. The skew is a measurement of a distribution's lack of symmetry and the kurtosis is a measurement of the peakedness (or pointiness) of a data distribution.

In general, if the values of either skewness or kurtosis are not close to zero, then the data set is not normally distributed. For an exhaustive discussion on the use of kurtosis, see (DeCarlo, 1997). Usually, a data transformation corrects skewness and/or kurtosis in the distribution of the original data.



a) b)

Figure 6.2: a) Skewness, and b) kurtosis

There are several common data transformations, including the *square-root transformation*: Here, the square root is taken of each data point/ observation. If the original data consists of negative numbers, a constant is added to each data point, e.g. $\sqrt{x_i + C}$ where $x_i$is the $i$th data point and $C$ is the added constant. If the range of data points are all measured in the range 0 to 1, then the *arcsine transformation* may be used, which consists of taking the arcsine of the square root of a number. The result is given in radians in the range from $-(\pi/2)$ to $(\pi/2)$. Another common data transformation is the log transform. In essence, the logarithm is taken from each data point. Usually, the base of the logarithm is immaterial, but most often $e$ (a.k.a Eulers constant, or the natural number) or 10 are used as the base. If the distribution is positively skewed (see the upper left sketch in figure 6.2), then in order to perform an analysis using parametric methods the original data values (i.e. each observation) $x$ can be $log10$-transformed into the transformed data values $x'$ as $x' = log10(x)$; where $x$ is the original data value and is the mean value of all the raters/assessors' individual estimates of the same image.

The $log10$-transformation makes a positively skewed data distribution less skewed. This operation was done to make patterns in the data more interpretable, as well as to accommodate the assumptions of parametric statistical tests. For more in-depth information about data transformations and their motivation, see for example, work by (Sokal and Rohlf, 2011), and (McDonald, 2014). Then after that, the $log10$-transformed data set, $x_i'$, can be normalised by subtracting the mean $log10$-value, $\overline{log10(x)}_i$ of all the human raters individual estimate of the same image $i$, as in equation 6.1:

$$Normalised(x_{ij}) = x_{ij}' - \overline{x'}_i \tag{6.1}$$

where $\overline{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_j$and each rater/assessor is denoted as $j$. This type of normalisation, which is a rescale-operation, is commonly named *per-example mean subtraction, mean-centring,* or *normalisation by subtracting the mean*. In essence it centres the distribution to have a zero mean value. This makes the distribution of the raters' estimate of each image more comparable with other images in the experiment.

To analyse the variance in between distributions (e.g. raters assessment of plant cover) and to fit linear models ANOVAs were conducted. After an ANOVA test has been conducted, it was followed by a residual analysis. A residual $\varepsilon_i$, is an estimate of the experimental error and is the difference between the observed value of the dependent variable $y_i$ and the predicted value $\hat{y}_i$ in the model, as in equation 6.2.

$$\varepsilon_i = y_i - \hat{y}_i \tag{6.2}$$

Where $i$ is the enumeration of all the observations, there is one residual $\varepsilon_i$ per observation $i$. The sum of all residuals and their mean is equal to zero. Residuals can be thought of as elements of variation that are not explained by the fitted model. The basic assumptions behind ANOVA and regression analysis are that the residuals are expected to be approximately normal and independently distributed, with a mean of 0 and some constant variance (Natrella et al., 2012). The residual analysis is usually carried out after the ANOVA to examine the fitted model.

## 6.3 Raters' Agreement/Reliability

Over the years several methods have been used to investigate whether the participating raters (also called assessors) are *in agreement with each other*, and to see how reliable their estimates are, Krippendorff summarised the essence of the many reliability measures as:

> *All reliability measures are intended to express the degree to which several assessors, several measuring instruments, or several interrogations of the same units of analysis yield the same descriptive accounts, category assignments, quantitative measures or data for short* (Krippendorff, 1992).

In this thesis, analysis-of-variance (ANOVA) tests were used to investigate whether if there were differences between the mean estimates reported by the raters. It tested the null hypothesis $H_0$ that the means of estimates are equal between the raters, see equation 6.3:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = ... = \mu_n \tag{6.3}$$

The alternative hypothesis $H_a$ is that at least two means are different from each other. The ANOVA is called an *omnibus* test and does not give information about which specific rater(s) were significantly different from each other. If this information is needed, then so called post-hoc tests are needed to determine which raters differed from each other. Examples of post-hoc tests include Tukey's honestly significant difference (Tukey's HSD) and the Bonferroni test. Further information about ANOVA and post-hoc tests can be found in (Sokal and Rohlf, 2011), (Field et al., 2012) and (Gardener, 2012).

It is also common to calculate the product-moment correlation in between two variables, e.g. Pearson's $\rho$ (or r) for parametric data (see equation 6.4), and similarly Spearman's $\rho$ and Kendalls $\tau$ for non-parametric data.

$$\rho_{X,Y} = \frac{E\left[(X-\mu_X)(Y-\mu_Y)\right]}{\sigma_X \sigma_Y} \tag{6.4}$$

where $X$ and $Y$ denotes the two datasets under investigation. $E$ is the expected value, $\mu$ denotes the mean of $X$ and $Y$, and $\sigma$ denotes the standard deviation of $X$ and $Y$ respectively.

The correlation coefficient measures linear agreement, i.e. whether if the measurements go up and down together (Dallal, 2012). The correlation coefficient is in the interval scale from -1 to +1, where +1 means perfect positive correlation (as one variable increases, the other variable will increase as well), and a zero value means no correlation at all (completely random association. For example, as one variable increases, the other variable sometimes increases, sometimes it decreases, or neither), and -1 means a perfect negative correlation (as one variable increases, the other variable decreases. So, it is an inverse relationship between the two variables). When it comes to investigating whether the *raters are in agreement*, it might not be enough to compute this coefficient. For example, if raters A and B assess four phenomena as A = (10, 20, 40, 50), respectively. B = (1, 2, 4, 5), then the correlation result ends up in a perfect positive correlation (r = 1). However, they are certainly not in agreement in their assessments.

A common test used to examine the reliability between two raters when using nominal (categorical) data is the Cohen $\kappa$-test (a.k.a. kappa test) (Cohen, 1960). This is commonly used in the fields of medicine and psychology, where, for example, two physicians (or psychologists) may assess a patient's symptoms. The outcome is a value called the $\kappa$-coefficient, which is in the interval between -1 and +1. A $\kappa$-value equal to +1 means perfect agreement between the two raters, and a $\kappa$-value of -1 implies perfect disagreement. If the $\kappa$-value equals zero, then there is no relationship between the ratings of the two raters. Hence, any agreement, or disagreement, occurs because of randomness. A $\kappa$-value, $\kappa = 0.7$ is generally considered to be a satisfiable level of agreement. There is also a *weighted $\kappa$-test* which is used for ordinal data. For further details, see work reported by (Cohen, 1968).

(Manel et al., 2001) compared the medical diagnostics and ecological modelling processes, and evaluated Cohen's $\kappa$-test as a way of assessing the absence, or presence of plant species. (Manel et al., 2001) pointed out the $\kappa$-coefficient as being meaningful when it comes to comparing models for predicting the distribution of organisms from environmental data.

Other reliability coefficients are, for example, the Scott's $\pi$ (pi)(Scott, 1955) for nominal data and two raters, the Fleiss' $\kappa$ (kappa) (Fleiss, 1971) also for nominal data but with a fixed number of several raters. Another, reliability coefficient is Krippendorff's $\alpha$ (alpha) which is of special interest since it applies to all levels of measurement (Krippendorff, 2004) and Hayes and Krippendorff (2007)) (see section 6.2). In this thesis raters typically estimated the extent of cover on an interval scale typically in between 0 to 100% (see section 4.3.1), or made frequency estimates using the binary absence/presence on a nominal scale. The Krippendorff's $\alpha$ is calculated using the general equation 6.5;

$$\alpha_{Kripp} = 1 - \frac{D_o}{D_e} \tag{6.5}$$

where $D_o$is the observed disagreements and $D_e$ is the expected random disagreements.

In order to assess reliability in terms of the consistency of measurements made by several raters measuring the same quantity, the ICC was calculated (Shrout and Fleiss, 1979). The ICC assesses the reliability of ratings by comparing the variability of different ratings of the same subject with the total variation across all ratings and all subjects.

There are several ICC classes; those that measure the *reliability of a single rater* are denoted as ICC(1,1), ICC(2,1), and ICC(3,1). In addition, there are classes that measure the *reliability of the mean rating*; these are denoted as ICC(1, k), ICC(2, k) and ICC(3, k). In this thesis the *reliability of a single rater* was used, i.e. the first three mentioned classes, above. In a study by (Shrout and Fleiss, 1979) these classes were described as:

*ICC1:* *Each target is rated by a different set of k judges, randomly selected from a larger population of judges.*

*ICC2:* *A random sample of k judges is selected from a larger population, and each judge rates each target, that is, each judge rates n targets altogether.*

*ICC3:* *Each target is rated by each of the same k judges, who are the only judges of interest.*

The raters participating in the investigations in this thesis were assumed to be representative of a larger number of similar raters in the population. Hence, the ICC(2,1) class was chosen (see definition in equation 6.6).

$$ICC(2,1) = \frac{var(\beta)}{var(\alpha) + var(\beta) + var(\varepsilon)} \tag{6.6}$$

where $var$ denotes the variance, $var(\alpha)$ denotes the variability due to differences in the rating scale used by the raters. For example, when considering a sample plot containing a ”true” value of 5% plant cover, rater A estimates this plot to contain 10% cover, but rater B estimates the same plot to contain 15%. Here, $var(\beta)$ denotes the variability caused by differences in the observed phenomenon/subjects (e.g. the sample plots containing plants), and $var(\varepsilon)$ denotes the variability caused by differences in the evaluations of the observed phenomenon/subjects by the judges. For example, rater A finds that a sample plot contains 45 plants, but rater D finds the same sample plot contains 5 plants, because of different personal opinions on what and how to count. The ICC(2,1) class is generalisable, whereas ICC values from ICC(3,1) class are not. The ICC coefficient can theoretically vary between 0 and 1.0, where an ICC value of 0 indicates no agreement (i.e,. no reliability), while an ICC value of 1.0 indicates perfect agreement/reliability (i.e,. the raters were unanimous in their decisions).

A Student’s (dependent) t-test can be used where there are two raters’ estimates. This test is used to compare two datasets when data in each sample set are related in a special way. For example, when a rater assesses the same phenomenon before and after (e.g. the plant density). In order to use this paired (or dependent) t-test (and not the independent t-test), the dataset must be organised in pairs, where there is a relationship between each pair of data points. The number of data points in each data set must be the same. When it comes to comparing two raters, or methods, it is worth noting that (Altman and Bland, 1983) and (Bland and Altman, 1986) criticised the use of (product-moment) correlation coefficients, difference between means (by using, for example, the paired t-test), and regression analysis methods for being inappropriate for analysing measurement methods. Qualitative ratings of ICC agreement based on the ICC values were suggested by (Cicchetti, 1994) as follows:

***Poor*** *ICC-value < 0.40*

***Fair*** *ICC-value 0.40 to 0.59*

***Good*** *ICC-value 0.60 to 0.74*

***Excellent*** *ICC-value 0.75 to 1.0*

For more detailed information on measuring raters' agreements and reliability using the aforementioned methods, see (Shoukri, 2003) and (Martin and Bateson, 2007). Details of how to use the ICC can be found in (Rankin and Stokes, 1998), (Hallgren, 2012) and (Weir, 2005). In this thesis ICC was calculated using the package *psych* in R (R Core Team, 2014).

## 6.4 Counting Plant Clusters/Patches

In order to be able to estimate the the number of individuals or plant patches, clusters, or tufts by counting it is important to know how they are defined. In these investigations, a plant cluster was defined as being an individual or tightly clumped group of individuals and was counted as one instance or occurrence. More precisely:

Let there be $n$ two-dimensional regions to represent clusters $c$ of vegetation. The clusters can be enumerated as $\sum_{i=1}^{n} c_i$, where each cluster has a geometric centre, $g_c$ (also called centroid). Each cluster $c$ is represented by an ellipse whose major axis, $a$, and minor axis, $b$, coincide with the Cartesian axes (see figure 6.3a). Assuming that the geometric centre $g_c$ has its coordinates at $(x_0, y_0)$ , then the Cartesian equation is:

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1 \tag{6.7}$$

The ellipse is used as a tool for roughly estimating a continuous line of vegetation patches, which forms the boundary of a closed geometrical figure, i.e., its perimeter (see figure 6.3b). By applying this operation, the centre of each patch can be computed using equation 6.7. Clusters can be joined to another identified cluster. Alternatively, clusters can be considered on their own, and defined as individual clusters in their own right. If an identified ellipse A has its centre of gravity intersecting another identified ellipse

B, then A and B are to be considered as one cluster represented by the area $A \cup B$ (see figure 6.4).

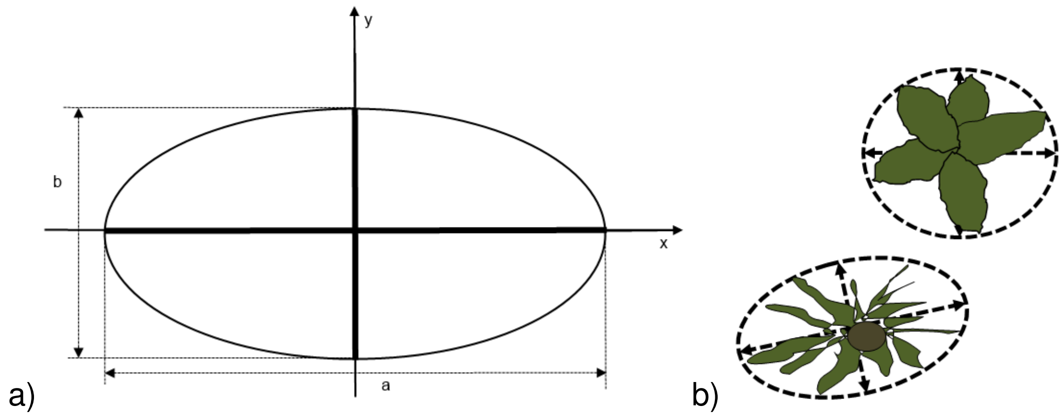



Figure 6.3: a) Ellipse, and b) Usage on plants/patches

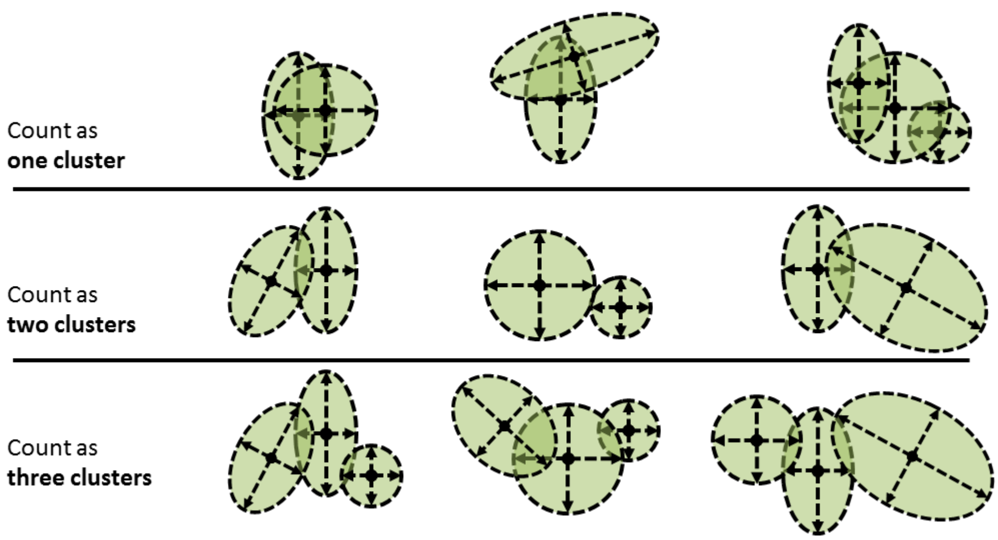



Figure 6.4: Interpretation of the definition of patches

The definitions in this section were used as a protocol for the raters assessing the number of individuals/clusters in the investigations in sections 6.8 , 6.9 and 6.11.

## 6.5 Six Studies in Reliability of Visual Estimates

Six studies were conducted to evaluate the reliability of (human) assessments made by two or more raters. The assessments refers to the estimates of vegetation extent, e.g. estimating the plant cover, or counting the number of plant individuals. The assessors in each investigation made their ratings/assessments from images or on-site.

In the following investigations presented in section: 6.6, 6.7, 6.8, 6.9, and 6.11, except for the 5$^{th}$ investigation (in section 6.10), parametric methods were used for the analysis of the current data setanalysis of variance (ANOVA) tests at 95% confidence level were carried out on the raters' visual estimates of total plant cover. The ANOVA tests were conducted on $\log_{10}$-transformed data. The purpose was to investigate differences between the raters' assessments.

## 6.6 1$^{st}$ Study of Visual Estimates of Images

Thirteen staff members at the Dalarna University, Sweden were picked at random and were asked to assess the total cover of vegetation in images. The underlying hypothesis was to test if many human raters are reliable in their estimates of cover. The raters neither possessed any experience in estimating plant cover nor did they have any experience of the railway domain.

### 6.6.1 Method

Images were acquired from a nadir view over the railway tracks of the Falun - Grycksbo railway (see details in section 5.3). Nine of those images were selected. Images showing totally overgrown track areas, as well as images that were relatively free of vegetation, were disregarded during this process.

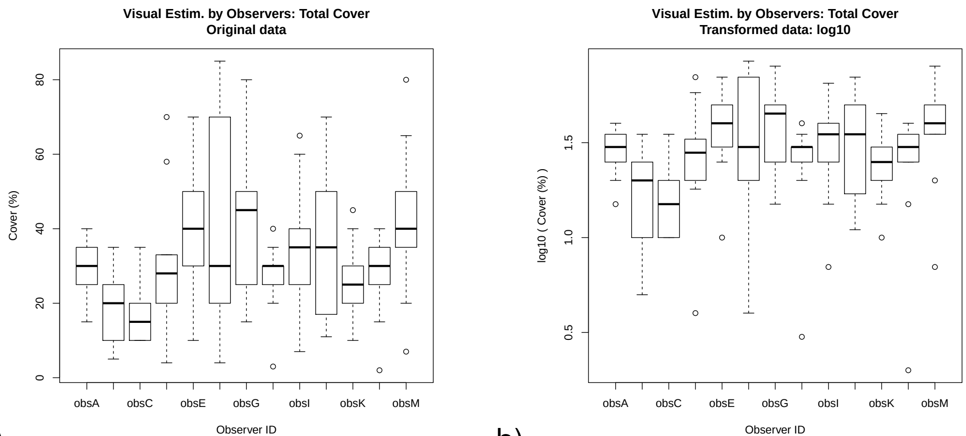
The raters were asked to individually conduct a visual estimate of the total plant cover in each of the nine images and report the same in terms of percentage cover relative to the image under observation. Only an area measuring 1 x 1 m (i.e., the sample plot), outlined on each image, was to be considered. The raters used an interval scale for estimating aerial cover and were free to choose any number between 0 and 100%. Only green parts of the plants (i.e., those containing chlorophyll) were to be included in the total cover. Each observation was made individually, with no interaction from other persons and no time constraints. Each rater wrote down their total cover estimate on a protocol template.

### 6.6.2 Results and Conclusions of the 1$^{st}$ Study

In total $n = 117$ visual estimates (observations) were made over all nine plots. Plot-wise observations reported by the 13 raters are presented in table C.1. These can be characterised with the following central tendencies concerning plant cover: $\overline{x} = 31.55\%$ and $Md = 30\%$ (see figure C.1a).

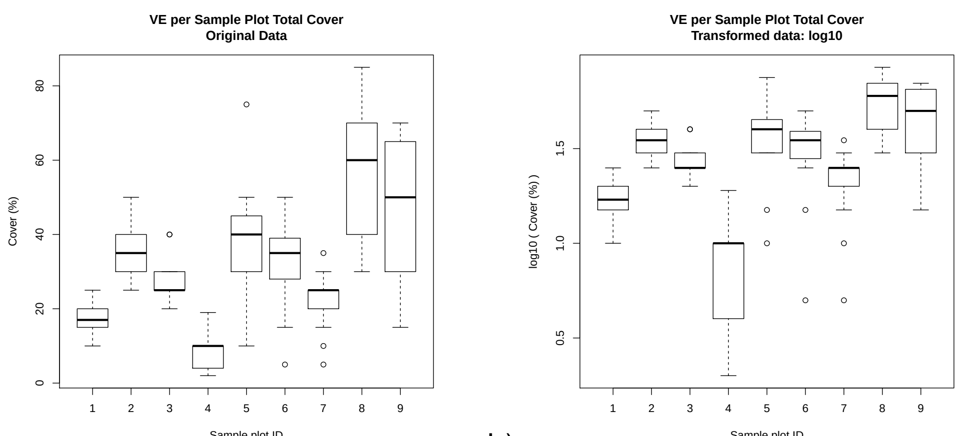
An ANOVA-test reported a *significant* difference in mean estimates. $F = 5.28$ at $df = 12\, and\, 104$, $p = 0.0000006896$. ($H_0$ was rejected at 0.05 significance level), i.e. the raters were not in agreement. (see figure 6.5).

The maximum difference when all the raters made a VE of the same plot was 65%, i.e., the highest estimate made by any rater minus the lowest estimate of any other rater of the same plot (see figure 6.6a). The differences between the maximum estimate and the minimum estimate per plot (from 1 to 9) were: 15, 25, 20, 17, 65, 45, 30, 55, 55%, respectively. These plot-wise differences in agreement can be summarised in the overall arithmetic mean, $\overline{x} = 36.33$, and the overall median, $Md = 30$. These central tendencies indicate high fluctuations, and emphasise the need for a strict protocol as well as the need for proper training before making visual estimates of plant cover.




a) b)

Figure 6.5: VE by each rater: a) original data, and b) log10-transformed data




b)

Figure 6.6: VE per sample plot: a) original data, and b) log10-transformed data

A measurement of reliability (or degree of agreement) was computed using the intraclass correlation coefficient, ICC(2,1) which concerned the level of agreement between the 13 raters when estimating total cover in nine images (see table 6.2). $ICC(2,1) = 0.61$, which may be referred to as a *moderate* agreement. This result, in conjunction with the ANOVA result, makes the reliability of visual estimates appear as weak.

| Estimate of | Reliability ICC(2,1) | no. of sample plots $n$ | p-value |
|---|---|---|---|
| Total Cover | $0.61$ | $9$ | $< 0.01$ |

Table 6.2: The levels of agreement among the raters

## 6.7 2$^{nd}$Study of Visual Estimates On-site

Three domain experts participated in an on-site investigation of the railway track in between Falun and Grycksbo (see section 5.3). The hypothesis was to test the reliability of humans in their estimates of cover.

### 6.7.1 Method

This investigation involved three raters with prior experience in estimating plant cover. In this context, it is worth mentioning that personnel working within the railway maintenance domain are not provided with any formal training in the assessment of the extent of vegetation. This is mostly because that vegetation management is scheduled as periodic maintenance (see chapter2). It is initiated by national railway authorities and is usually sub-contracted to other companies.

A person in charge of vegetation inspections was asked to select two representative railway segments (strata) that characterise different levels of (mostly) herbaceous vegetation. Only track segments hosting vegetation were considered. The two strata were each about 100 m long and were classified as: low level coverage, and high level of coverage, respectively. In each stratum, five sample plot positions were randomised by simple random (see figure 6.7). Each sample plot had an area of 1 x 1 m. All sample plots were digitally photographed from a nadir view.

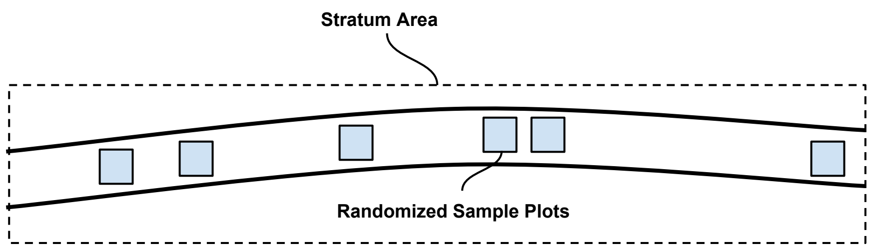



Figure 6.7: Randomised plots in a stratum

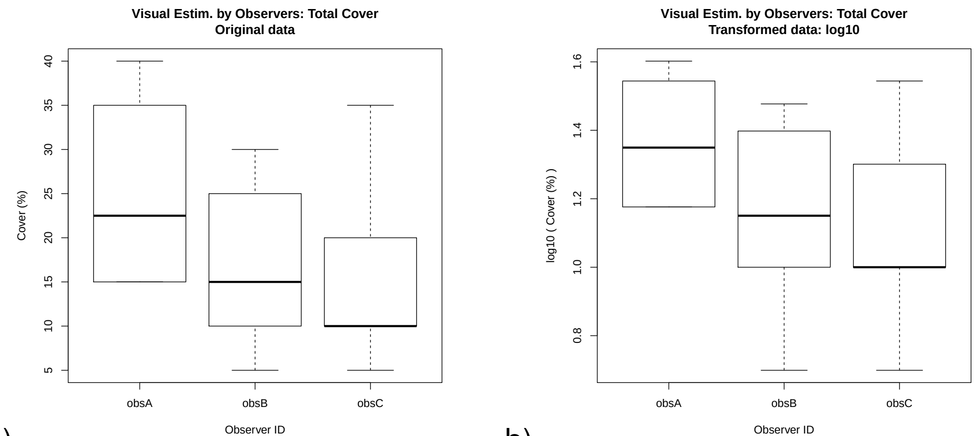
The raters were asked to separately report a visual estimate of the total plant cover in each of the 10 sample plots. The raters were instructed to estimate the total vegetation cover (from 0 to 100% to the nearest 5%) in each sample plot area. No further instructions of how to make their judgements were given. Details about the data collected could be found in section 5.3.

### 6.7.2 Results and Conclusions of the 2nd Study

In all $n = 30$ visual estimates (observations) for assessing total plant cover were made. Plot-wise observations reported by the three raters are presented in table C.2. Central tendencies for all observations: $\overline{x} = 18.83\%$ and $Md = 17.5\%$ (see figure C.3a).

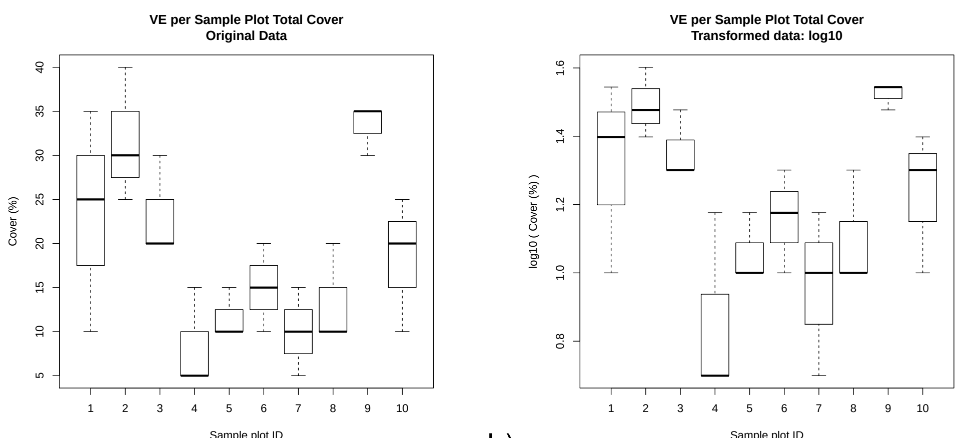
An ANOVA-test reported a *significant* difference in mean estimates was reported. $F = 19.72$ at $df = 2\, and\, 27$, $p = 0.000005265$. ($H_0$ was rejected at 0.05 significance level), i.e. they were not in agreement (see figure 6.8).

The maximum difference when all the raters made a VE of the same plot was 25%, i.e., the highest estimate made by any rater minus the lowest estimate of any other rater of the same plot (see figure 6.9a). Differences between the maximum estimate and the minimum estimate per plot (from 1 to 10) were: 25, 15, 10, 10, 5, 10, 10, 10, 5, 15%, respectively. These plot-wise differences in agreement can be summarised in the overall arithmetic mean, $\overline{x} = 11.5$, and the overall median, $Md = 10$.




a) b)

Figure 6.8: VE by each rater: a) original data, and b) log10-transformed data




b)

Figure 6.9: VE per sample plot: a) original data, and b) log10-transformed data

To assess the reliability (or degree of agreement) between the three raters when estimating the total cover in 10 sample plots, the intraclass correlation coefficient, ICC(2,1), see table 6.3. As in the investigation in section 6.6, the results of this investigation show similarities. The level of agreement, $ICC(2,1) = 0.56$, in this investigation when using domain experts may be referred to as *moderate*. Again, the latter result in conjunction with the ANOVA result makes the reliability of visual estimates appear as weak.

The results support the hypothesis that humans are unreliable in their estimates of cover extent.

| Estimate of | Reliability ICC(2,1) | no. of sample plots $n$ | p-value |
|---|---|---|---|
| Total Cover | $0.56$ | 10 | $3.5 * 10^{-5}$ |

Table 6.3: The three raters' levels of agreement

## 6.8 3rdStudy of Visual Estimates of Images

Three raters (laypersons) were asked to assess the cover extent and the sub-plot frequency on images. This investigation was a follow up of the investigation made in section 6.6. Based on the results of the investigations made in sections 6.6 and 6.7, more guidance was given to the raters so that they knew what and how to make visual estimates.

The first hypothesis was that laypersons are not in agreement with each other when it comes to estimating the extent of cover and sub-plot frequency. The second hypothesis was that laypersons would benefit from having been presented with a rigorous protocol before making the assessments, both in terms of the estimating of cover, and the counting of plant clusters. This protocol is described in the next section.

### 6.8.1 Method

Three raters were picked at random from among the academic staff at Dalarna University, Sweden. The raters did not possess any experience in estimating plant cover; nor did they have any experience of the railway domain. For each presented image, each rater was asked to:

1. estimate the cover of woody plants (in %)
2. estimate the cover of herbs (in %)

3. estimate the cover of grass (in %)

This was done three times using the following rater methods:

1. Aerial canopy cover (ACC)
2. Aerial foliage cover (AFC)
3. Sub-plot frequency

After each rater had finished making VEs of the attributes, as shown in the list above, they were also asked to count the number of vegetation clusters in each image. This procedure of counting was carried out twice: once after they made their VE using AFC and once after using ACC.

**Plant cluster definition**: In this investigation, a plant cluster was defined as being an individual plant or a tightly clumped group of individual plants. The protocol of how to count clusters presented in section 6.4. This was presented to the assessors.

| Estimate of/Method | Aerial Canopy Cover | Aerial Foliage Cover | Sub-plot frequency |
|---|---|---|---|
| Woody plants | **35** | **35** | **35** |
| Herbs | **35** | **35** | **35** |
| Grass | **35** | **35** | **35** |
| Counting Plant Patches | **35** | **35** | **-** |

Table 6.4: Overview - number of observations per person

Note that the raters were not informed that the same images were shown each time and in the same sequence. All raters were instructed on how assess cover using the three methods. They were also given advance assistance in how to estimate cover (see figure 6.10). They were not allowed to use this assistance during the assessments. The purpose of showing them the visual cover aid chart was to visually synchronise (or symbolically calibrate) their feeling of how much area coverage is required to meet a certain percentage.

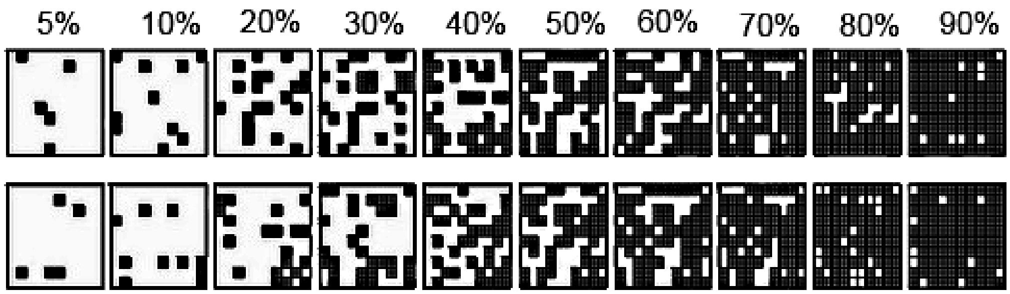



Figure 6.10: Visual aid chart for estimating cover

Since it is relatively easy to assess whether a plot contains nothing or is completely full, images showing either 0% and 100% of the attribute in question were removed before the analysis.

### 6.8.2 Visual Estimates of Plant Cover and Sub-plot Frequency

Nine investigations were carried out. As stated, three raters participated. Each rater visually assessed the cover extent or sub-plot frequency (SF) in a maximum of 35 images using three different methods, ACC, AFC, and SF, respectively. The target plants were woody plants, herbs, and grass. The images used for SF contained a 10 x 10 sub-plot sampling frame (see figure 6.11 as an example). The other images for the assessments of ACC and AFC did not contain a grid.

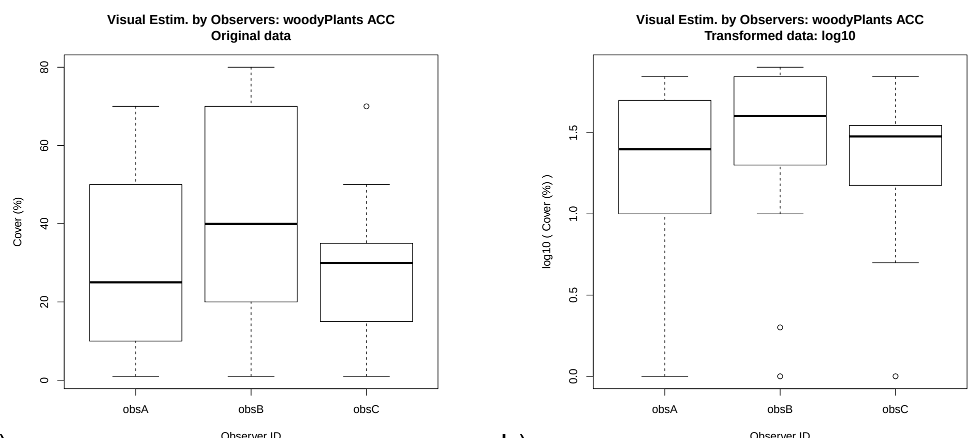
Figure 6.11: Sample area with a 10 x 10 sub-plot frame

### VE of woody plants using ACC method

The number of observed images in the analysis was 21 out of 35. The rest of the images were not selected because they did not contain any woody plants. The distribution of the data is presented in figure C.5

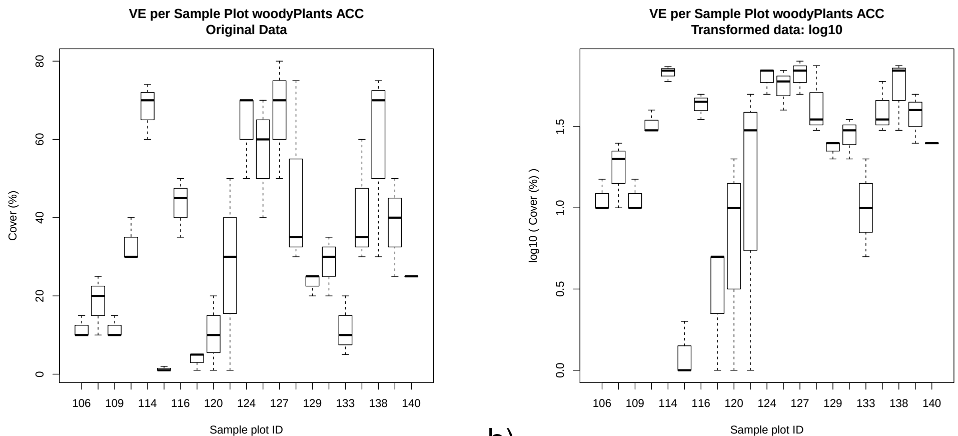
**Results**: An ANOVA test reported a *significant* difference in mean estimates. $F = 4.943$ at $df = 2$ and $60$, $p = 0.0103$ ($H_0$ was rejected at 0.05 significance level), i.e. the raters were not in agreement (see the box-plot in figure 6.12b).

Plot-wise variation and differences are presented in figure 6.13. The maximum difference when the raters made an VE of the same plot was 49%, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: Md = 15% and the mean: $\overline{x} = 18.9\%$.




a) b)

Figure 6.12: VE by each rater: a) original data, and b) log10-transformed data




b)

Figure 6.13: VE per sample plot: a) original data, and b) log10-transformed data

### Visual Estimates of Woody Plants Using the AFC Method

The number of observed images in the analysis was 22 out of a total of 35. The remaining images were excluded due to the lack woody plants.

**Results**: An ANOVA test reported a *significant* difference in mean estimates. $F = 10.5$ at $df = 2$ and $63$, $p = 0.0001158$. ($H_0$ was rejected at 0.05 significance level), i.e. the raters were not in agreement (see boxplot in figure 6.14b).

Plot-wise variation and differences are presented in figure 6.15. The maximum difference when the raters made a VE of the same plot was $35\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was $0\%$, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimates: $Md = 6\%$ and the mean: $\overline{x} = 9.0\%$.

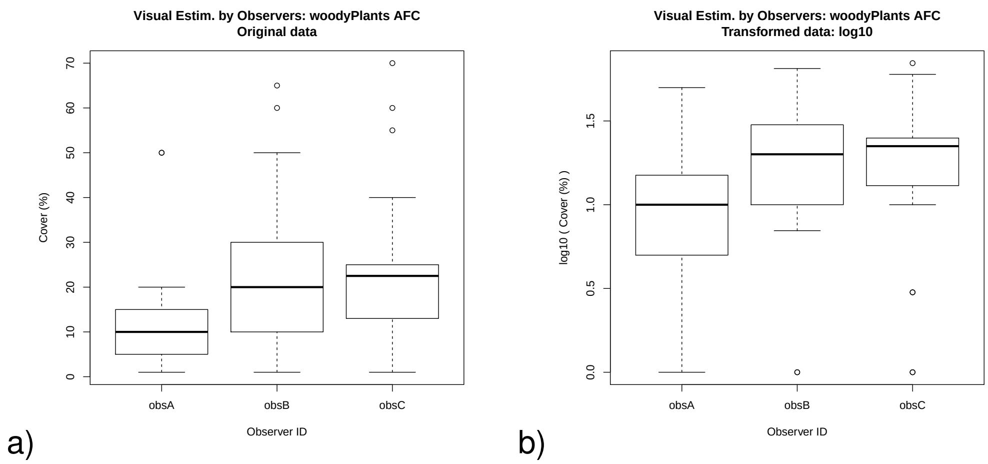



Figure 6.14: VE by each rater: a) original data, and b) log10-transformed data

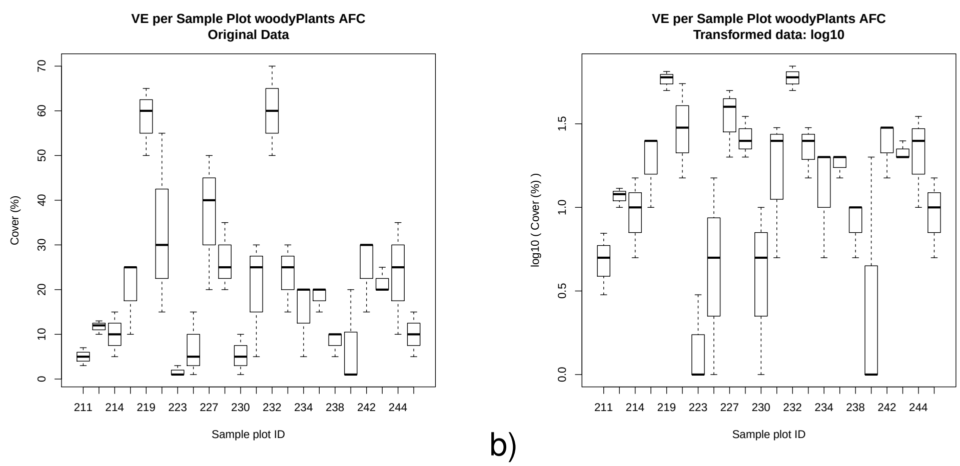



Figure 6.15: VE per sample plot: a) original data, and b) log10-transformed data

### VE of woody plants using SF method

The number of observed images in the analysis was 23 out of 35. The rest of the images were not selected because they did not contain any woody plants.

**Results**: An ANOVA test reported a *significant* difference in mean estimates. $F = 9.897$ at$df = 2$ and $66$, $p = 0.0001741$. ($H_0$ was rejected at 0.05 significance level) (see the box-plot in figure 6.16b).

Plot-wise variation and differences are presented in figure 6.17. The maximum difference when the raters made a VE of the same plot was $75\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 14\%$ and the mean: $\overline{x} = 20.0\%$.

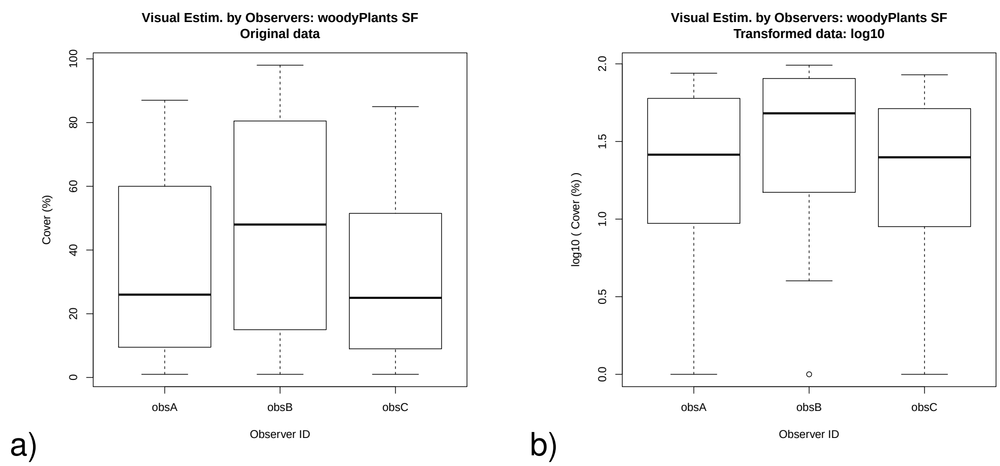
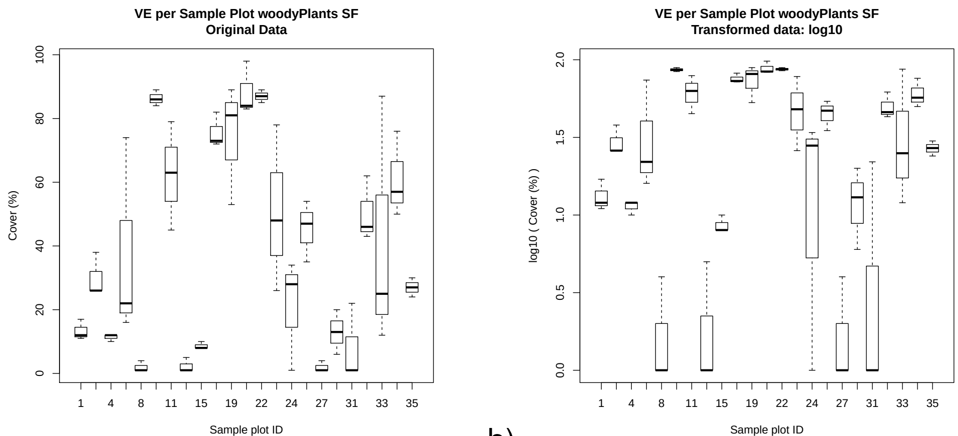



Figure 6.16: VE by each rater: a) original data, and b) log10-transformed data




b)

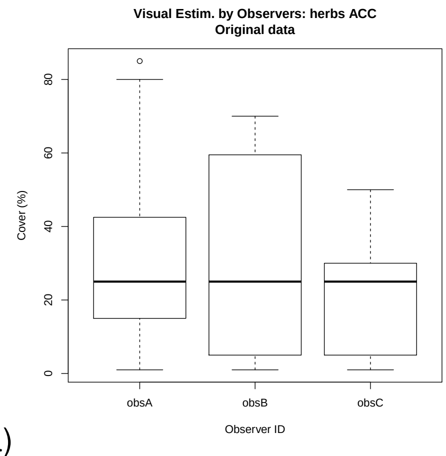
Figure 6.17: VE per sample plot: a) original data, and b) log10- transformed data

### Visual Estimates of Herbs using the ACC Method

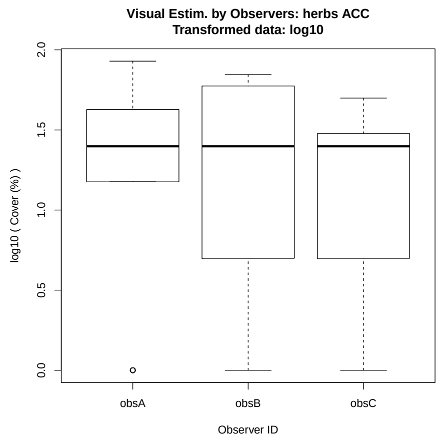
The number of observed images in the analysis was 23 out of 35. In cases where all raters unanimously estimated 0% cover of the target plant, these images were removed.

**Results:** An ANOVA test reported a *non-significant* difference in mean estimates. $F = 1.391$ at $df = 2$ and $66$, $p = 0.256$. ($H_0$ was *not* rejected at 0.05 significance level) (see the box-plot in figure 6.18b).

Plot-wise variation and differences are presented in figure 6.19. The maximum difference when the raters made a VE of the same plot was $79\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was $0\%$, i.e. total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate: Md = 19% and the mean: $\overline{x} = 22.43\%$.

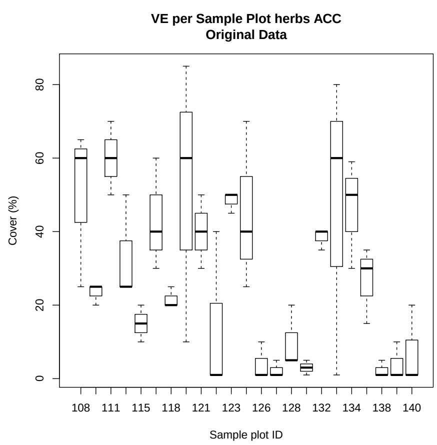


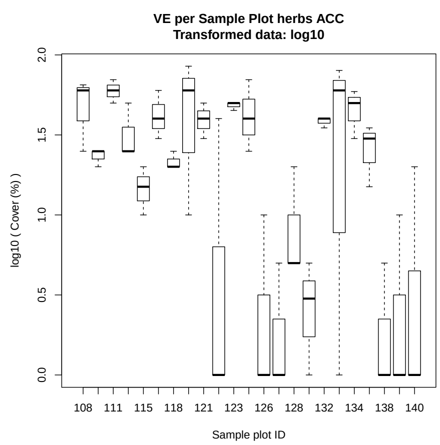

a)




b)

Figure 6.18: VE by each rater: a) original data, and b) log10- transformed data




b)

Figure 6.19: VE per sample plot: a) original data b) log10-transformed data

### Visual Estimates of Herbs Using the AFC Method

The number of observed images in the analysis was 21 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

**Results:** An ANOVA test reported a *significant* difference in mean estimates. $F = 3.955$ at $df = 2$ and $60$, $p = 0.02435$. ($H_0$ was rejected at 0.05 significance level) (see the boxplot in figure 6.20b).

Plot-wise variation and differences are presented in figure 6.21. The maximum difference when the raters made a VE of the same plot was $65$%, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 15$% and the mean: $\overline{x} = 16.4$%.

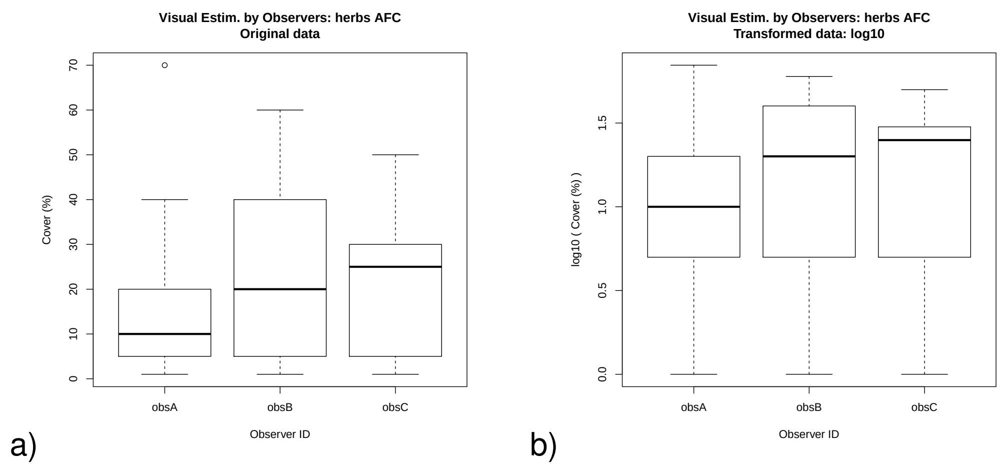
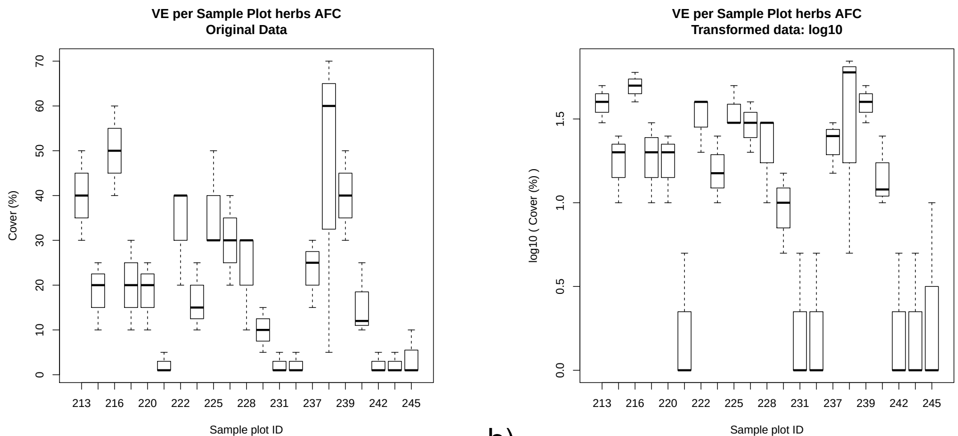



Figure 6.20: VE by each rater: a) original data, and b) log10- transformed data




b)

Figure 6.21: VE per sample plot: a) original data, and b) log10- transformed data

### VE of Herbs using SF method

The number of observed images in the analysis was 21 out of 35. The rest of the images were not selected because they did not contain any herbs.

**Results**: An ANOVA test reported a *non-significant* difference in mean estimates. $F = 2.129$ at $df = 2$ and $60$, $p = 0.1278$. ($H_0$ was *not* rejected at 0.05 significance level) (see the box-plot in figure 6.22b). Plot-wise variation and differences are presented in figure 6.23. The maximum difference when the raters made a VE of the same plot was $85$%, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md =$ 17% and the mean: $\overline{x} = 25.1\%$.

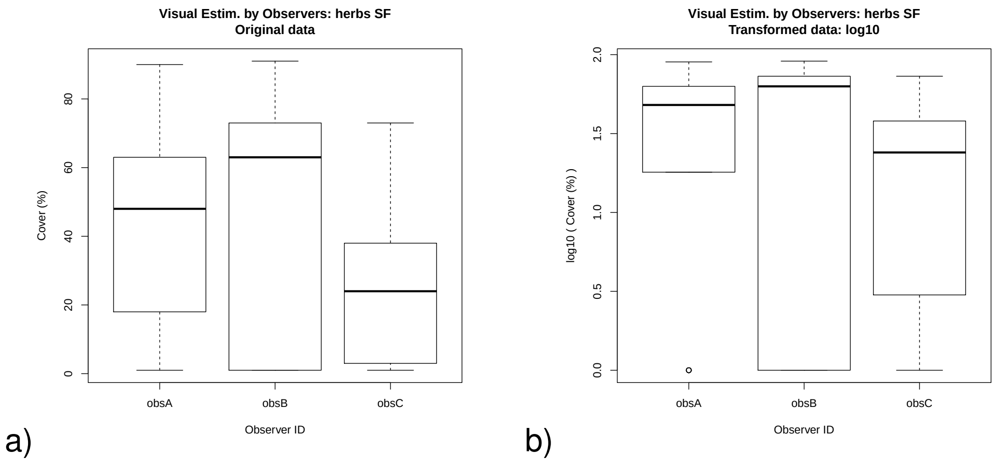



Figure 6.22: VE by each rater: a) original data, and b) log10-transformed data

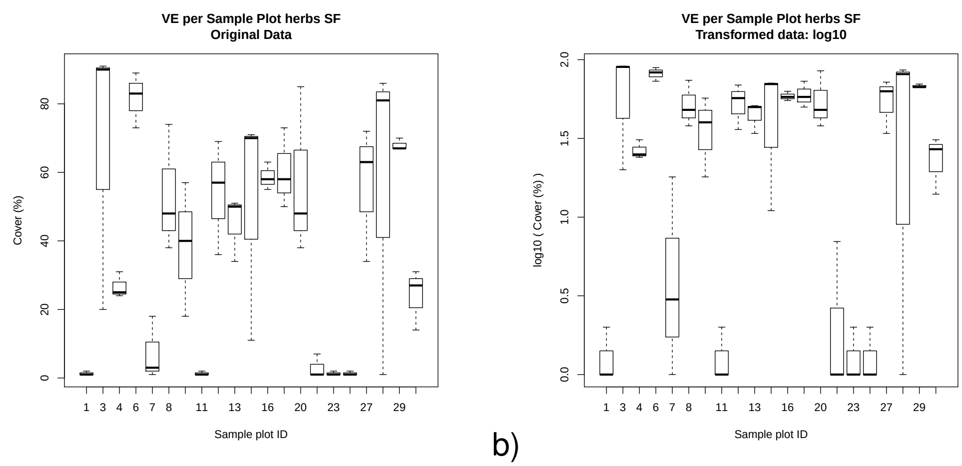
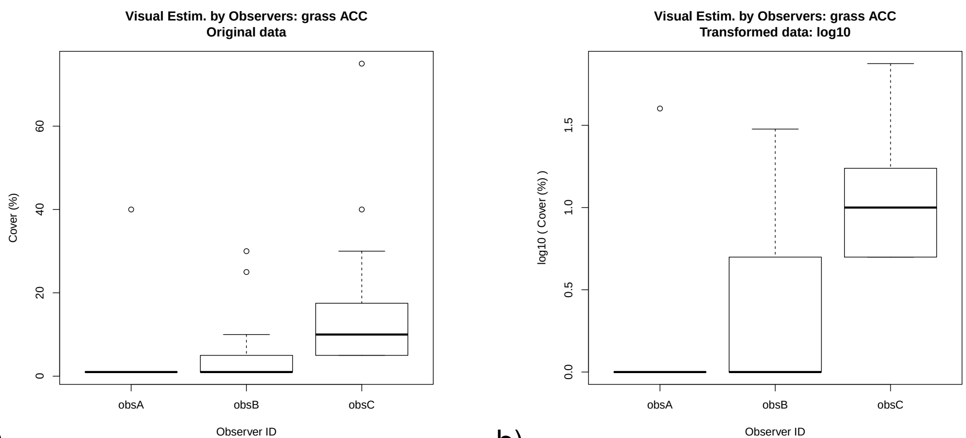



Figure 6.23: VE per sample plot: a) original data, and b) log10-transformed data

### Visual Estimates of Grass Using the ACC Method

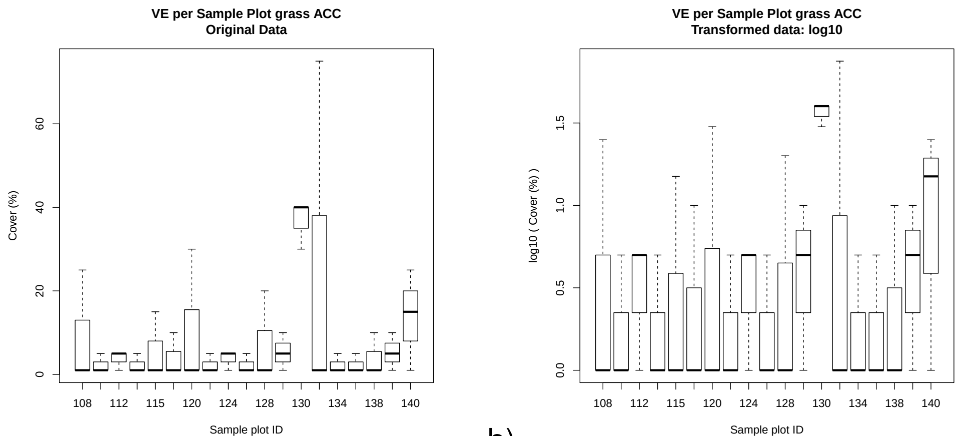
The number of observed images in the analysis was 19 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

**Results:** An ANOVA test reported a *significant* difference in mean estimates. $F = 55.62$ at $df = 2$ and $54$, $p = 7.676 * 10^{-14}$. ($H_0$ was rejected at 0.05 significance level), see the box-plot in figure 6.24b. Plot-wise variation and differences are presented in figure 6.25. The maximum difference when the raters made a VE of the same plot was $74$%, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 9$% and the mean: $\bar{x} = 13.9$%.




a) b)

Figure 6.24: VE by each rater: a) original data, and b) log10- transformed data




b)

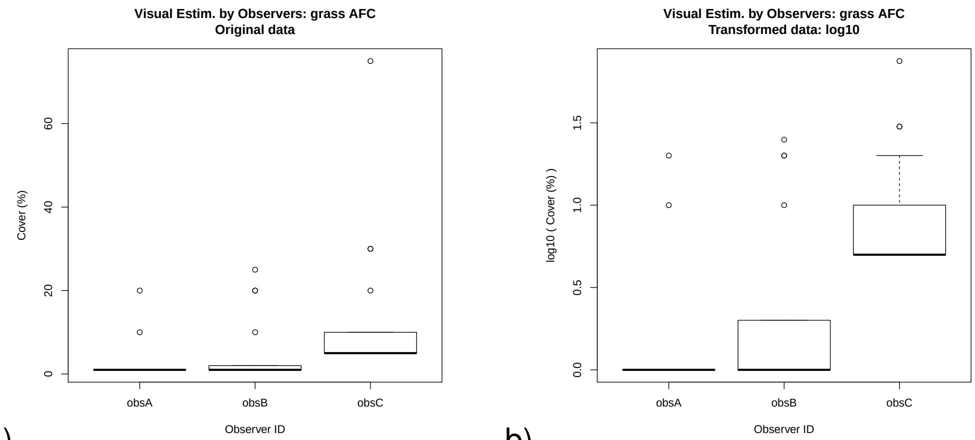
Figure 6.25: VE per sample plot: a) original data, and b) log10- transformed data

### Visual Estimates of Grass using the AFC Method

The number of observed images in the analysis was 19 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

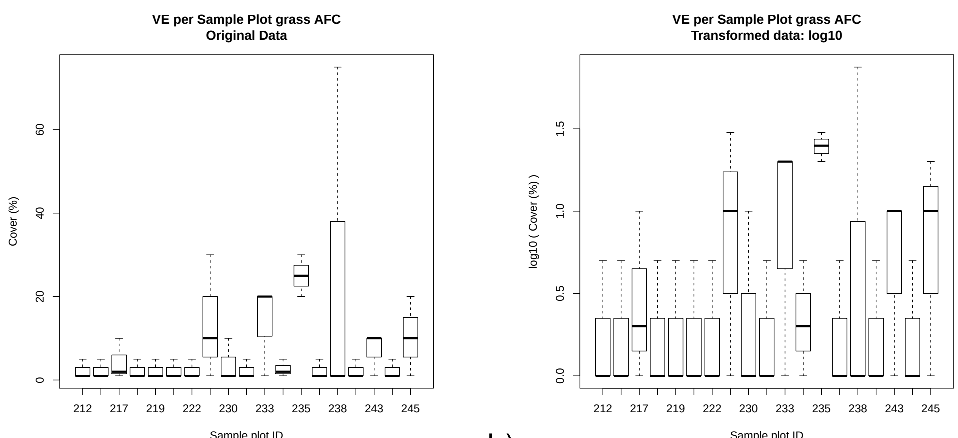
**Results:** An ANOVA test reported a *significant* difference in mean estimates. $F = 48.94$ at $df = 2$ and $54$, $p = 7.472 * 10^{-13}$. ($H_0$ was rejected at 0.05 significance level) (see the box-plot in figure 6.26b).

Plot-wise variation and differences are presented in figure 6.27. The maximum difference when the raters made a VE of the same plot was $74$%, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 4$% and the mean: $\bar{x} = 11.6$.




a) b)

Figure 6.26: VE by each rater: a) original data b) log10-transformed data




b)

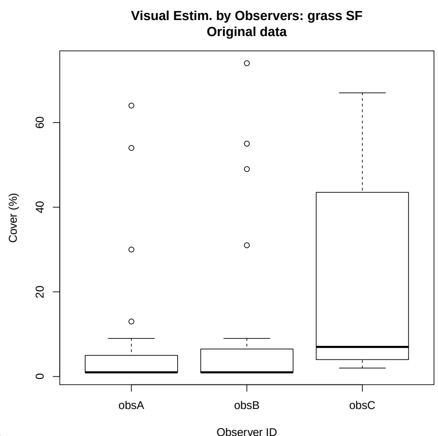
Figure 6.27: VE per sample plot: a) original data, and b) log10-transformed data

### Visual Estimates of Grass using the SF method

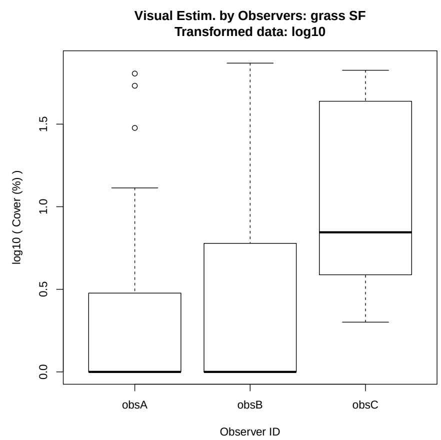
The number of observed images in the analysis was 20 out of 35. The rest of the images were not selected because they did not contain any grass.

**Results**: An ANOVA test reported a *significant* difference in estimates was reported. $F = 21.21$ at $df = 2$ and $57$, $p = 0.0000001304$. ($H_0$ was rejected at 0.05 significance level), see the box-plot in figure 6.28b.

Plot-wise variation and differences are presented in figure 6.29. The maximum difference when the raters made an VE of the same plot was $66\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 6.5\%$ and the mean: $\bar{x} = 20.15\%$.

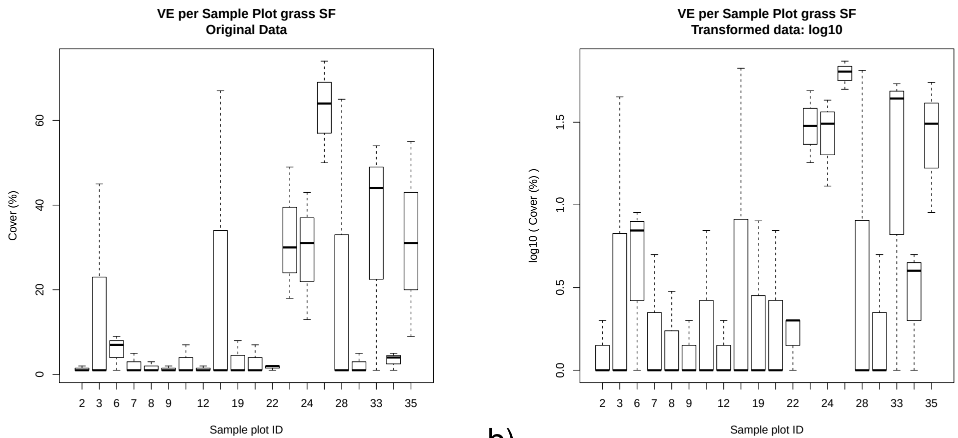


a)



b)

Figure 6.28: VE by each rater: a) original data, and b) log10-transformed data




b)

Figure 6.29: VE by each rater: a) original data, and b) log10- transformed data

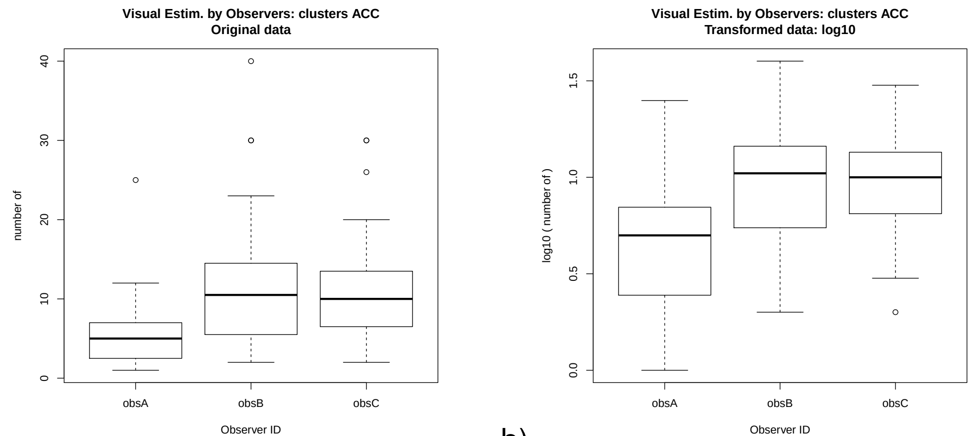
## 6.8.3 Counting Plant Clusters

After estimating cover using the AFC and AFC methods, the participants were asked to count the number of plant clusters in the images. All three raters were presented with the definition of a plant cluster as defined in 6.4.

### Inter-rater Agreements in Counting Plant Clusters

Inter-rater agreement is the degree of agreement among the three raters when estimating the number of plant clusters from the same set of images.

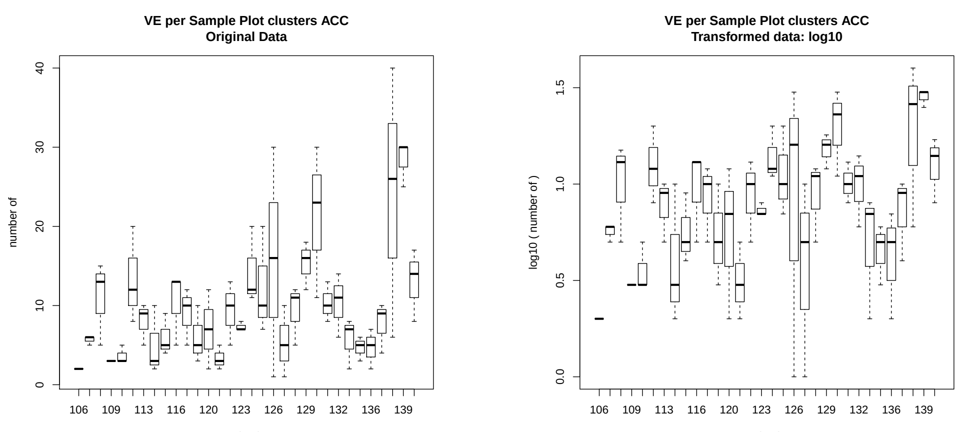
**Results:** An ANOVA test was performed on the raters' counting of plant clusters after the ACC session. A *significant* difference in count estimates between the raters was reported . $F = 58.08$ at$df = 2$ and $93$, $p < 2.2 * 10^{-16}$. ($H_0$ was rejected at 0.05 significance level) (See box-plot in figure 6.30b).




a) b)

Figure 6.30: Each rater's counting of plant clusters after the ACC session: a) original data, and b) log10-transformed data

Plot-wise variation and differences are presented in figure 6.31. The maximum difference between the raters in counting the same plot was $34$ plant clusters, i.e., the highest count minus the lowest count in the same plot (see figure 6.31). The median difference (over all plots) between the highest and lowest plot was: $Md = 7$ plant clusters and the mean: $\overline{x} = 8.12$ plant clusters.




b)

Figure 6.31: Plant cluster counts per sample plot: a) original data, and b) log10-transformed data

Another ANOVA test, also at 95% confidence level, was performed, but this time on the raters' counting of plant clusters after the AFC session. A *significant* difference in count estimates between the raters was also reported. $F = 66.43$ at $df = 2$ and $93$, $p < 2.2 * 10^{-16}$. ($H_0$ was rejected at 0.05 significance level) (see box-plot in figure 6.32b).

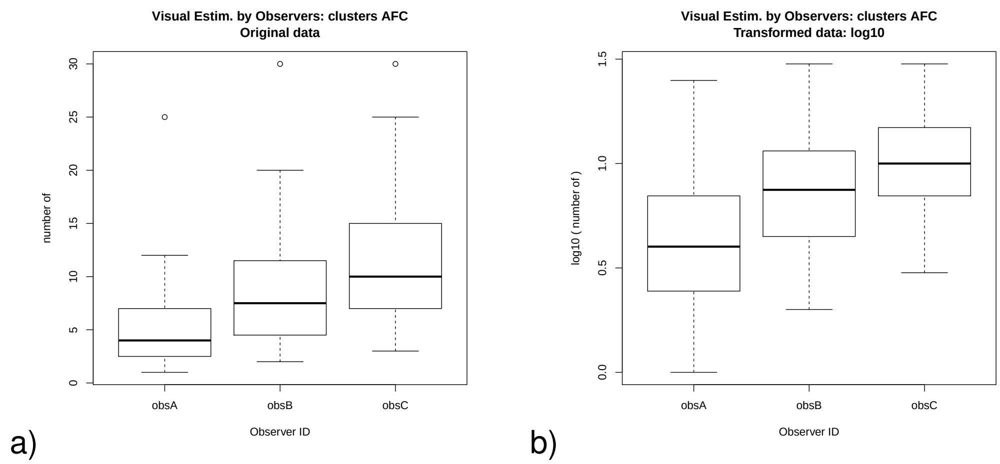
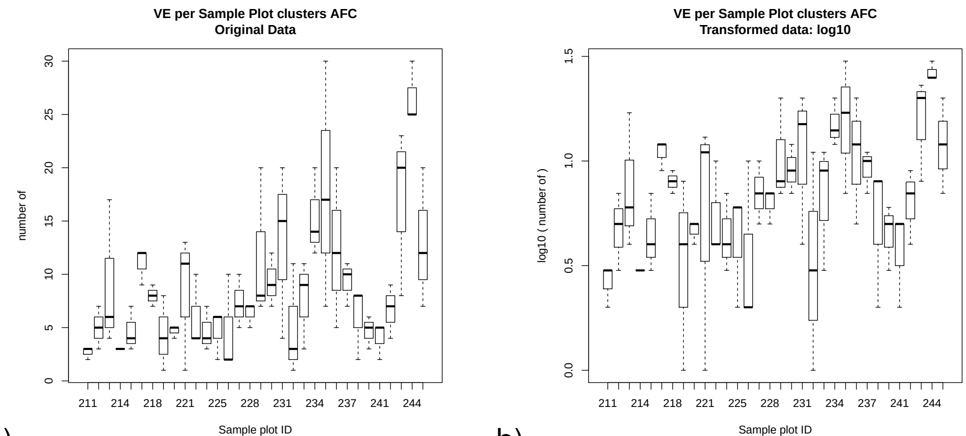



Figure 6.32: Each rater's counting of plant clusters after the AFC session: a) original data, and b) log10-transformed data

Plot-wise variation and differences are presented in figure 6.33. The maximum difference between the raters in counting the same plot was 23 plant clusters, i.e., the highest count minus the lowest count in the same plot (see figure 6.33). The median difference (over all plots) between the highest and lowest plot was: Md = 5 plant clusters and the mean: $\overline{x} = 7.12$ plant clusters.



a) b)

Figure 6.33: Plant cluster counts per sample plot: a) original data, and b) log10-transformed data

To assess the extent of agreement between the three raters, intraclass correlation coefficients ICC(2,1) were calculated for the two sessions of counting of plant clusters (see table 6.5).

| Count during session | Reliability ICC(2,1) | no. of images $n$ | p-value |
|---|---|---|---|
| ACC | $0.45$ | $32$ | $4.5 * 10^{-10}$ |
| AFC | $0.41$ | $32$ | $1.3 * 10^{-9}$ |

Table 6.5: ICC(2,1) for agreement between the three raters' counting

Also the Krippendorff's $\alpha$ was also calculated: ACC counting session $\alpha_{Kripp} = 0.463$, and for the AFC counting session $\alpha_{Kripp} = 0.478$

### Intra-rater Agreements in Counting Plant Clusters

The intraclass analysis was conducted by computing several pairwise student's t-tests. Two sets of paired samples were collected from each rater. The estimates were made by the same rater on the same images, first during the ACC session and again during the AFC session.

The test results on the pairs' transformed log10-data is presented in table 6.6.

For each table row, 34 degrees of freedom applies, df = n − 1 where the number of images n = 35. The mean difference in estimates between raters A and C was non-significant, $p > 0.05$, but rater B's mean difference in estimates was significantly different, $p < 0.05$

| rater ID | Counting after the ACC session | Counting after the AFC session | Mean of the differences, $\overline{D}$ | t(34) | p-value |
|---|---|---|---|---|---|
| A | $\bar{x} = 0.581$, $SE = 0.059$ | $\bar{x} = 0.543$, $SE = 0.058$ | 0.038 | 1.184 | 0.2447 |
| B | $\bar{x} = 0.897$, $SE = 0.070$ | $\bar{x} = 0.788$, $SE = 0.062$ | 0.109 | 4.065 | 0.0003 |
| C | $\bar{x} = 0.899$, $SE = 0.064$ | $\bar{x} = 0.912$, $SE = 0.063$ | −0.013 | −0.688 | 0.4962 |

Table 6.6: The raters' two (paired) sessions of counting plant clusters, df = 34

### 6.8.4 Conclusion

**Visual estimates of plant cover:** Between the three raters, seven out of the nine ANOVA tests resulted in significant differences in the mean estimates of cover ($p < 0.05$) (see sections 6.8.2 to 6.8.2 for details. Only two classes (herbs, ACC, and herbs SF) were found to be non-significant, meaning that no difference could be observed between the raters in those two cases. However, some of the raters occasionally found it hard to differentiate between a pine and a tuft of grass, when seen in nadir perspective. This could explain some of the fluctuations in cover estimates.

The raters' levels of agreement were assessed by computing the intracorrelation coefficient ICC(2,1). The degrees of freedom used for the calculations in table 6.7 are $df1 = (n-1)$ and $df2 = (o-1)(n-1)$, where $o$ is the number of raters, and $n$ is the number of estimated images.

| Target plants | VE method | Reliability ICC(2,1) | p-value | Details in section |
|---|---|---|---|---|
| Woody Plants | ACC | 0.68 | $1.5 * 10^{-8}$ | 6.8.2 |
| | AFC | 0.62 | $3.9 * 10^{-8}$ | 6.8.2 |
| | SF | 0.76 | $1.1 * 10^{-12}$ | 6.8.2 |
| Herbs | ACC | 0.58 | $2.4 * 10^{-6}$ | 6.8.2 |
| | AFC | 0.73 | $7.4 * 10^{-10}$ | 6.8.2 |
| | SF | 0.78 | $4.1 * 10^{-11}$ | 6.8.2 |
| Grass | ACC | 0.15 | 0.011 | 6.8.2 |
| | AFC | 0.19 | 0.003 | 6.8.2 |
| | SF | 0.43 | $2.4 * 10^{-5}$ | 6.8.2 |

Table 6.7: The three raters' level of agreement as of the ICC(2,1)

The ICC(2,1) results shown in 6.7 give values between $ICC(2,1) = 0.15$ (for the estimate of grass cover, ACC) up to $ICC(2,1) = 0.78$ (for the estimate of herb cover using SF). Using the suggested qualitative ratings put forward by (Cicchetti, 1994), the obtained results were characterised as being *poor agreement* up to *good agreement* between the raters.

**Counting plant clusters**

*Inter-rater reliability*: The raters were asked to count the number of clusters in all images during the ACC and AFC sessions. ANOVA tests showed that there were *significant differences* between the raters counting of plant clusters. In addition, ICC values were calculated (see table 6.5), and the values were found to be similar to when the raters were estimating the cover extent (see table 6.7): $ICC(2,1) = 0.45$ and $ICC(2,1) = 0.41$ respectively. In both cases, the result points to the lower boundary of a *moderate* agreement between the raters in counting plant clusters. Similar results were obtained by computing Krippendorff's $\alpha$, which also characterises the results as *moderate* agreement (ACC$\alpha_{Kripp} = 0.463$ and AFC $\alpha_{Kripp} = 0.478$).

*Intra-rater reliability*: Mean differences between, first, the ACC, and, second, the AFC VE sessions for each rater (A to C) were compared using paired student's t-tests. Raters A and C showed individual stability in counting when the same test was repeated twice, resulting in *non-significant* differences for the mean (p>0.05). The mean of differences together with each session standard errors were low (see table 6.6). Rater B had more difficulties in counting the same images twice, which resulted in a *significant* mean difference ($p<0.05$).

## 6.9 4th Study of Visual Estimates of Images

Five maintenance engineer administrators representing all four national territories of Swedish national railway administration, namely STA North, STA South, STA East, STA West, and Borlange municipality were asked to make visual estimates from 51 images that showed the railway embankment from a nadir (bird's-eye) perspective. Common to these maintenance engineer administrators was that each one was ultimately responsible for determining whether vegetation management should be carried out, or not, within their territory, or municipality.

### 6.9.1 Method

This investigation was carried out as an online Internet survey. The respondents (i.e., the raters) from the STA encompassed the full population of maintenance engineer administrators in Sweden. First, the respondents were contacted by telephone, individually. Then, each respondent was instructed through slide-show presentations and web conferencing software as to what to do and how to make the VE. Each respondent was able to browse the images from a website. The respondents had no contact with each other.

All the participating raters were instructed on where to make their assessment, i.e., defining the sampling area (see the yellow sampling area in figure 6.34a), and how to make their assessments using the AFC and ACC methods (see figure 6.34b). A review of these two methods of observation can be found in section 4.3.1. Before each rater made a series of VEs using AFC, or ACC, the order of the images was randomly re-arranged.

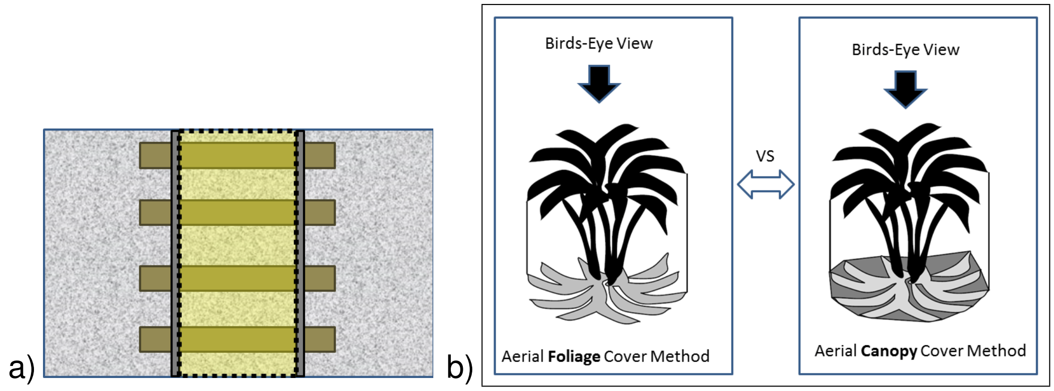



Figure 6.34: a) Sampling area, and b) AFC vs ACC method

By using these methods, the raters were asked to make VE of the cover on a scale of 0 to 100%, concerning:

1. Woody plants
2. Herbs
3. Grass
4. The remainder, i.e. everything else e.g. gravel, soil, wood, rocks etc, also including litter, such as dead plants.

After finishing making VEs of the attributes shown in the list above, the raters were also asked to count the number of vegetation clusters in each image. This procedure of counting was carried out twice, once after they made their VE using AFC, and once after using ACC.

**Plant cluster definition**:In this investigation, a plant cluster was defined as being an individual plant or a tightly clumped group of individual plants. These were counted, as shown in section 6.4.

| Estimate of/Method | Aerial Canopy Cover | Aerial Foliage Cover |
|---|---|---|
| Woody plants | 51 | 51 |
| Herbs | 51 | 51 |
| Grass | 51 | 51 |
| Counting Plant Patches | 51 | 51 |

Table 6.8: Overview - number of observations per person

Since it is relatively easy to assess whether a plot contains nothing, or is completely full of the plant type in question, images with 0% and 100% of the attribute in question were removed before the analysis. The purpose of this investigation was to see whether or not the raters would give the same estimate (i.e., a difference in estimates). It was not in the interests of this work to try to find out which raters differed in their estimates. Therefore no post-hoc tests were performed.

### 6.9.2 Visual Estimates of Plant Cover

Six investigations were carried out. Each rater visually assessed the cover extent in 51 images using two different methods, ACC and AFC, respectively. The target plants were woody plants, herbs, and grass.

The original data set was $log10$-transformed, $x_i'$, and then normalised by subtracting the mean $log10$-value, $\overline{log10(x)}_i$ of all the human raters individual estimate of the same image $i$, as in equation 6.1. This operation centres the distribution to have a zero mean value, and makes the distribution of the raters' estimate of each image more comparable with other images.

#### Visual Estimates of Woody Plants Using the ACC Method

The number of observed images in the analysis was 42 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

**Results**: An ANOVA test reported a *non-significant* difference in estimates was reported. $F = 1.499$ at $df = 4$ and $205$, $p = 0.2037$. ($H_0$ could not be rejected at 0.05 significance level) (see6.35).

Plot-wise variation and differences are presented in figure 6.36. The maximum difference when the raters made a VE of the same plot was $79\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was $0\%$, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 14\%$ and the mean: $\overline{x} = 19.6\%$.

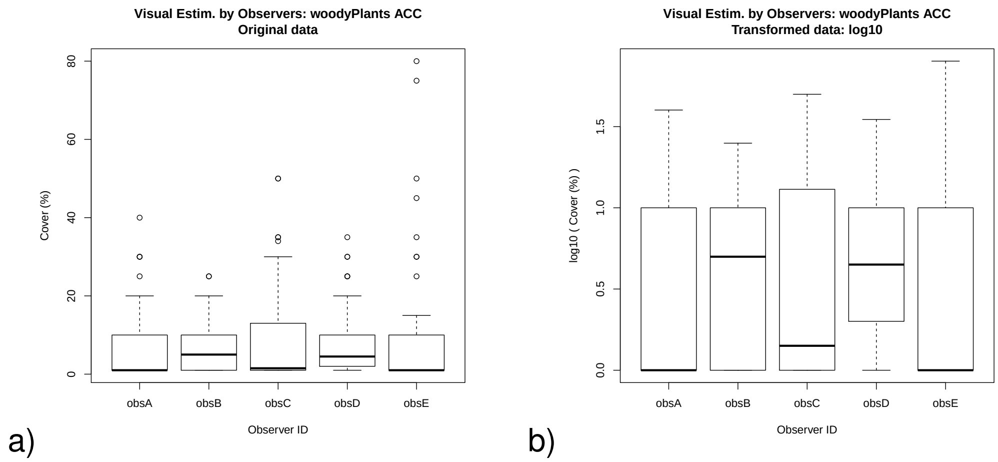



Figure 6.35: VE by each rater: a) original data, and b) log10- transformed data

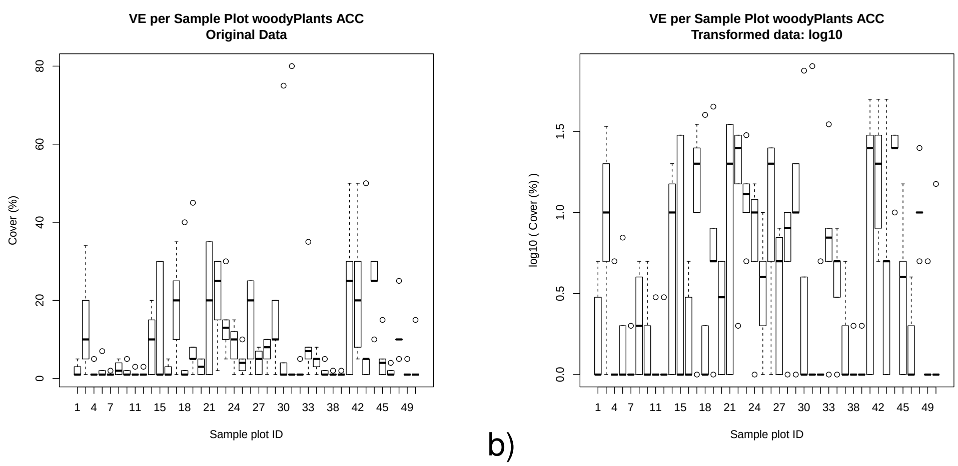



Figure 6.36: VE per sample plot: a) original data b) log10-transformed data

### Visual Estimates of Woody Plants Using the AFC Method

The number of observed images in the analysis was 38 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

**Results**: An ANOVA test reported a *significant* difference in estimates was reported. $F = 5.219$ at $df = 4$ and $185$, $p = 0.0005262$. ($H_0$ was rejected at 0.05 significance level) (see box-plots in figure 6.37b).

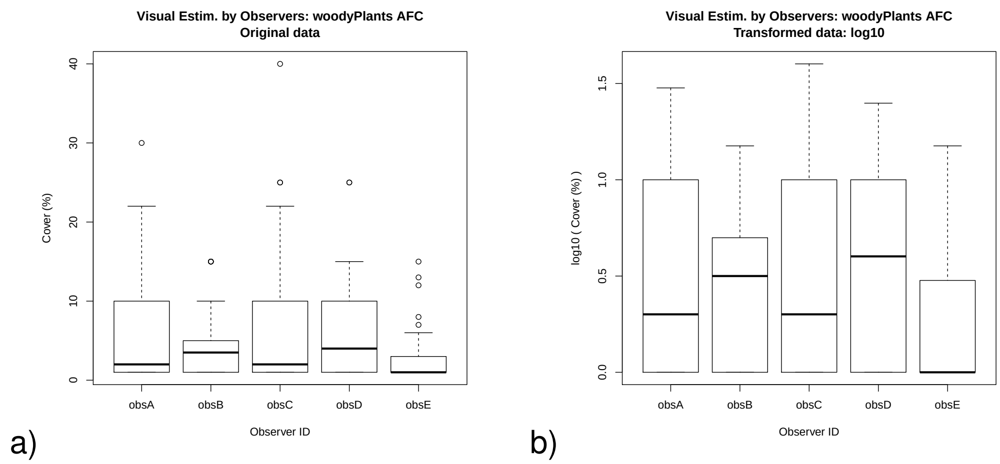
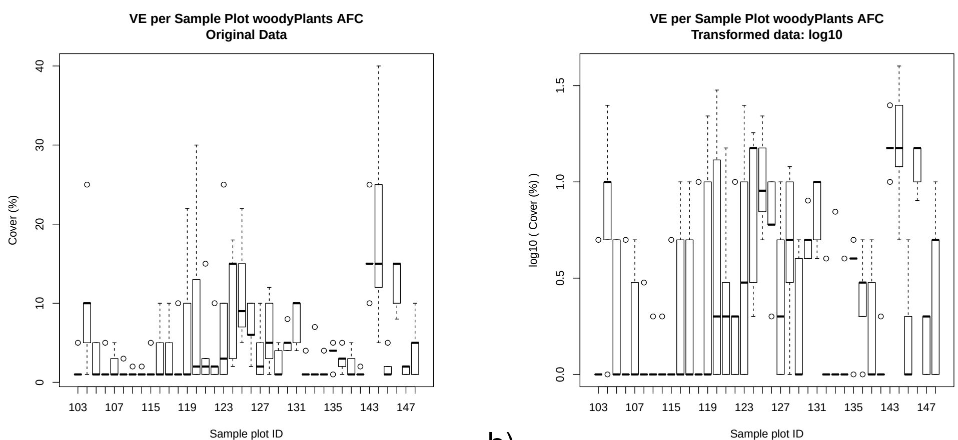



Figure 6.37: VE by each rater: a) original data b) log10-transformed data

Plot-wise variation and differences are presented in figure 6.38. The maximum difference when the raters made an VE of the same plot was $35\%$, i.e. the highest estimate minus the lowest estimate in the same plot. The minimum difference was $0\%$, i.e. total agreement. The median difference (over all plots) between highest and lowest plot cover estimate: $Md = 6\%$ and the mean: $\overline{x} = 9.0\%$.




b)

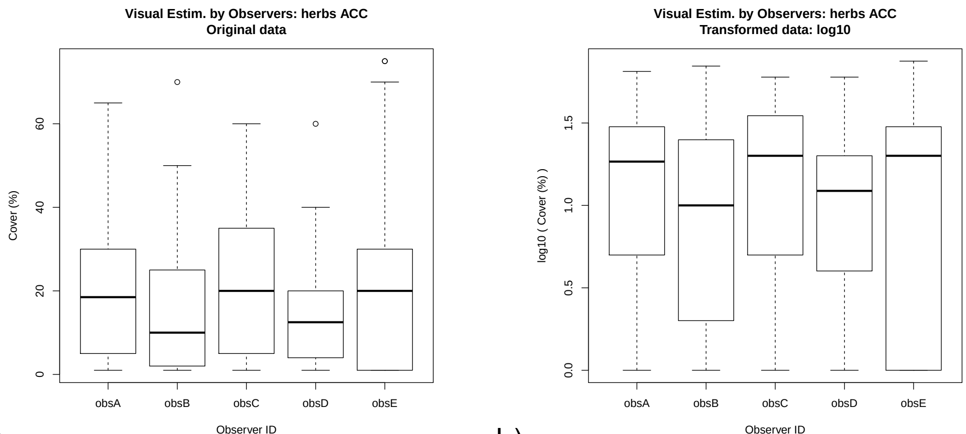
Figure 6.38: VE per sample plot: a) original data b) log10-transformed data

### Visual Estimates of Herbs Using the ACC Method

The number of observed images in the analysis was 42 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

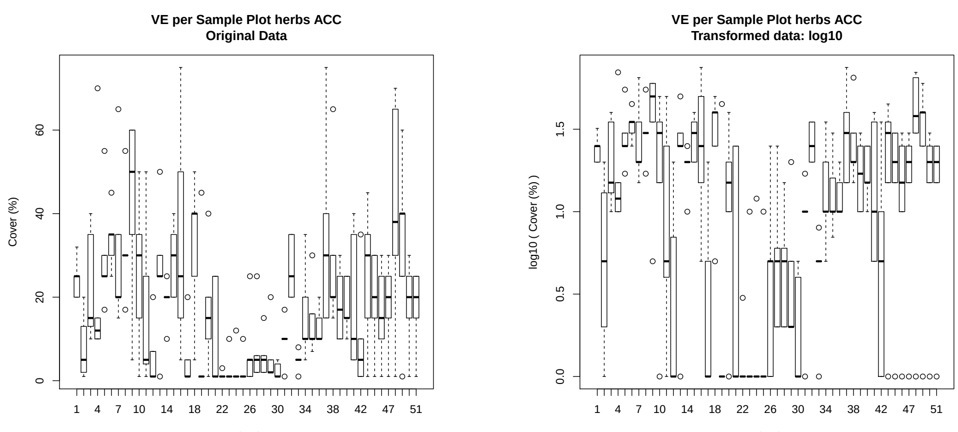
**Results:** An ANOVA test reported a *non-significant* difference in estimates was reported. $F = 1.227$ at $df = 4$ and $245$, $p = 0.2998$. ($H_0$ could not be rejected at 0.05 significance level), see 6.39).

Plot-wise variation and differences are presented in figure 6.40. The maximum difference when the raters made a VE of the same plot was $70\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was 0%, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate:$Md = 26.5\%$ and the mean: $\overline{x}$ = 30.0%.




a) b)

Figure 6.39: VE by each rater: a) original data b) log10-transformed data




b)

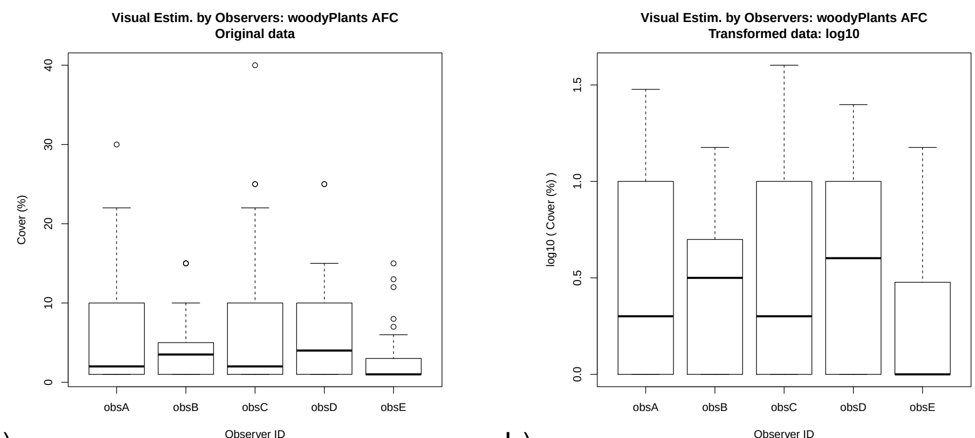
Figure 6.40: VE per sample plot: a) original data, and b) log10-transformed data

### Visual Estimates of Herbs Using the AFC Method

The number of observed images in the analysis was 47 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

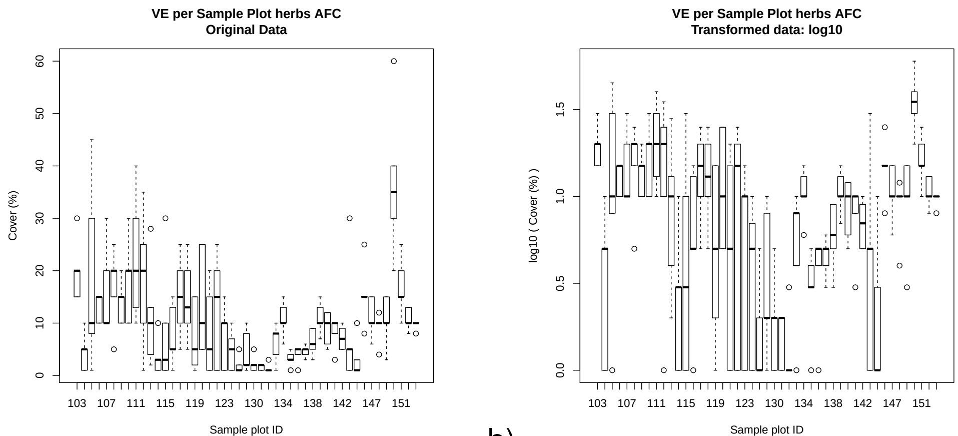
**Results:** An ANOVA test reported a *significant* difference in estimates was reported. $F = 4.674$ at $df = 4$ and $230$, $p = 0.001206$. ($H_0$ was rejected at 0.05 significance level) (see section 6.41).

Plot-wise variation and differences are presented in figure 6.42. The maximum difference when the raters made a VE of the same plot was $44\%$, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was 0%, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 9\%$ and the mean: $\bar{x} = 13.9\%$.




a) b)

Figure 6.41: VE by each rater: a) original data, and b) log10-transformed data




b)

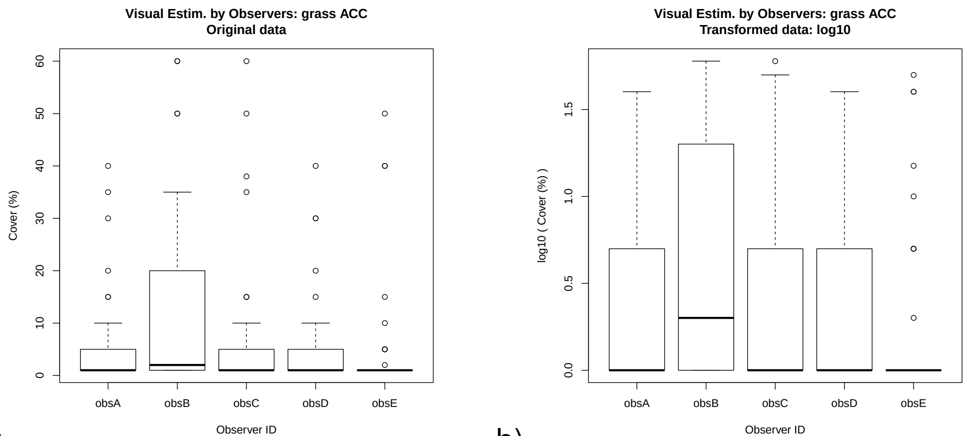
Figure 6.42: VE per sample plot: a) original data b) log10-transformed data

### Visual Estimates of Grass Using the ACC Method

The number of observed images in the analysis was 37 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

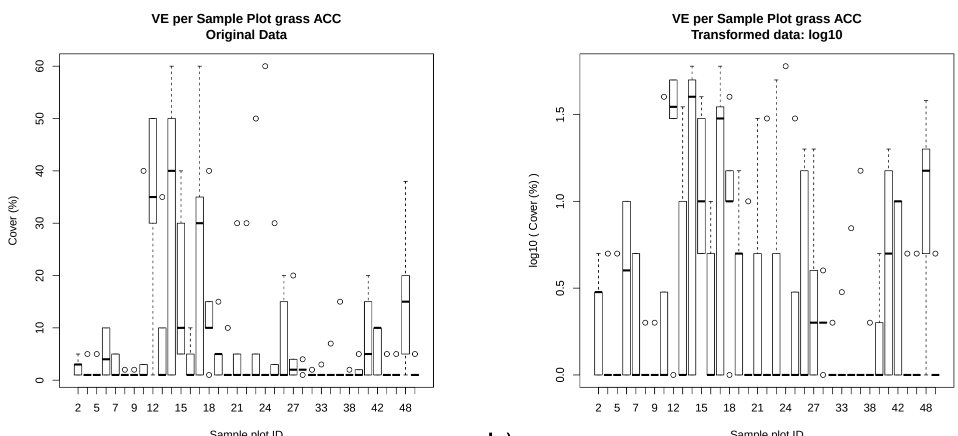
**Results:** An ANOVA test reported a *significant* difference in estimates was reported. $F = 2.666$ at $df = 4$ and $180$, $p = 0.03395$. ($H_0$ was rejected at 0.05 significance level) (see 6.43).

Plot-wise variation and differences are presented in figure 6.44. The maximum difference when the raters made a VE of the same plot was $59$%, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was 0%, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 9$% and the mean: $\overline{x} = 19.3$%.




a) b)

Figure 6.43: VE by each rater: a) original data, and b) log10-transformed data




b)

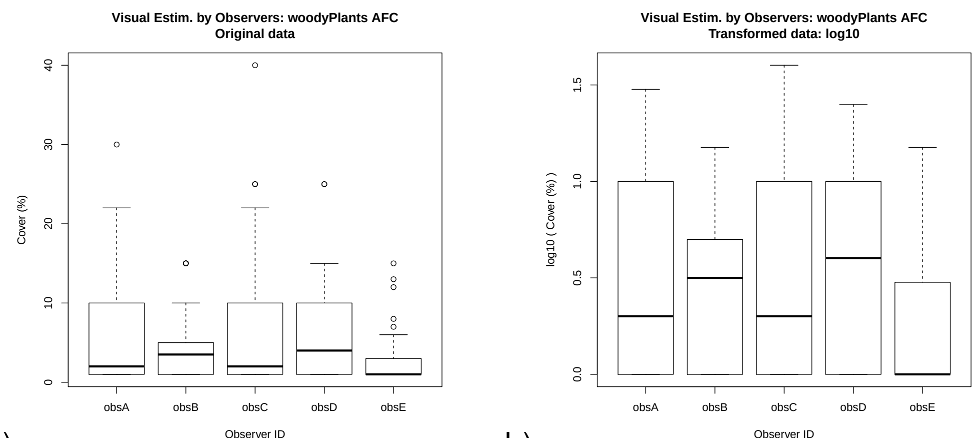
Figure 6.44: VE per sample plot: a) original data, and b) log10-transformed data

### Visual Estimates of Grass using the AFC Method

The number of observed images in the analysis was 33 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

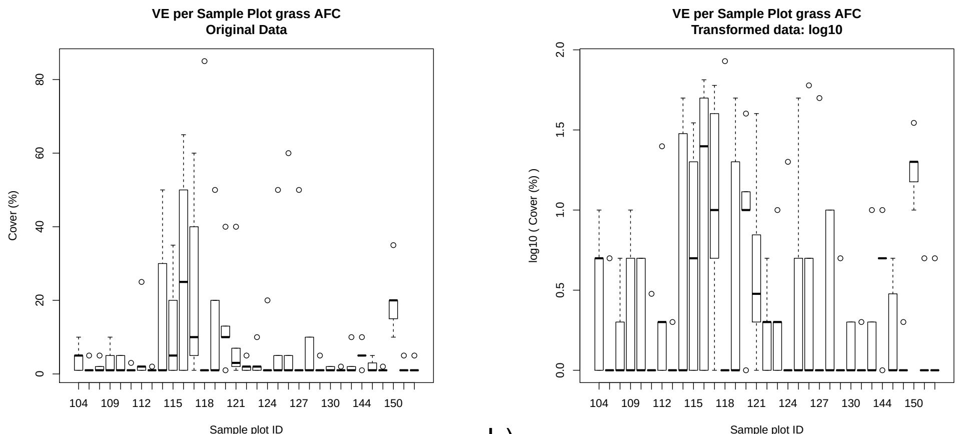
**Results:** An ANOVA test reported a *significant* difference in estimates was reported. $F = 5.427$ at $df = 4$ and 160, $p = 0.0004012$. ($H_0$ was rejected at 0.05 significance level) (see figure 6.45).

Plot-wise variation and differences are presented in figure 6.46. The maximum difference when the raters made a VE of the same plot was 84%, i.e., the highest estimate minus the lowest estimate in the same plot. The minimum difference was 0%, i.e., total agreement. The median difference (over all plots) between the highest and lowest plot cover estimate: $Md = 9\%$ and the mean: $\bar{x} = 22.2\%$.




a) b)

Figure 6.45: VE by each rater: a) original data, and b) log10-transformed data




b)

Figure 6.46: VE per sample plot: a) original data, and b) log10-transformed data

### 6.9.3 Counting Plant Clusters

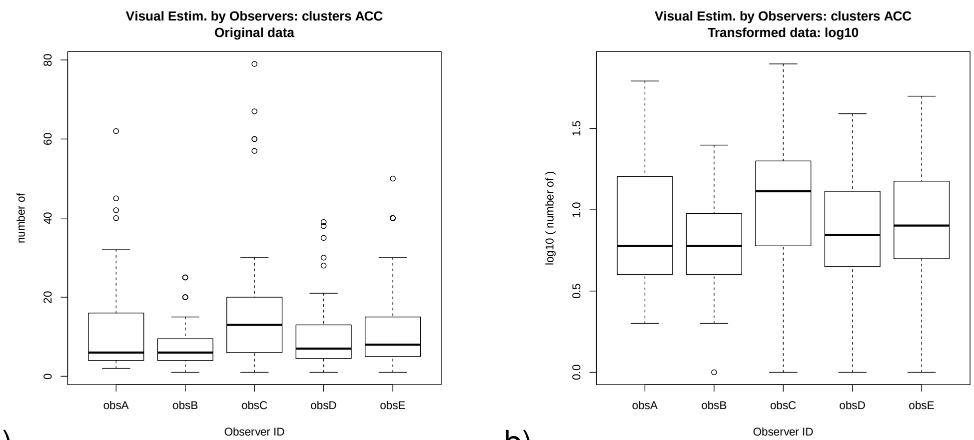
In this investigation, the raters had to estimate the number of plant clusters (as defined in section 6.4, by counting them in each image, $n = 51$. For each rater, the counting occurred twice using the same images at both occasions. The first time counting session is denoted as ”after the ACC session”, or just ACC, and the second time counting session is denoted as ”after the AFC session”, or just AFC.

While the original data was not normally distributed, the original data points $x_i$ were transformed into $\hat{x} = log10(x_i), \sum_{i=0}^{n=51} x_i$ to be able to use parametric methods for the data analysis (both for the inter-rater agreements and intra-rater agreements). Concerning on the raters’ counting of plant clusters, all ANOVA-tests were conducted at 95% confidence level.

#### Inter-rater Agreements in Counting Plant Clusters

Inter-rater agreement , (or inter-rater reliability) is the degree of agreement between the five raters who estimated the number of plant clusters from the same set of images. The term ”inter-rater” implies ”between raters”.

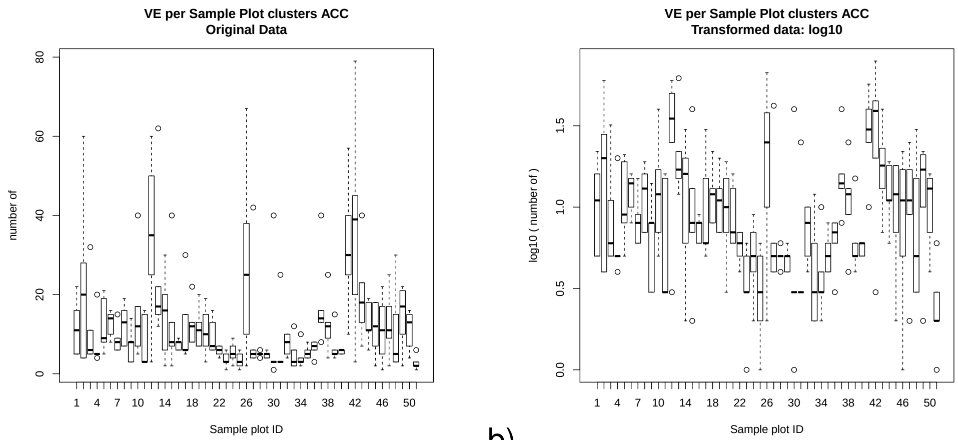
**Results of the raters' 1st counting of plant clusters:** After the ACC-session, an ANOVA-test was conducted on the counting results. It reported a *significant* difference in count estimates between the raters was reported. $F = 5.579$ at $df = 4\,and\,250$, $p = 0.0002566$. ($H_0$ was rejected at 0.05 significance level) (see figure 6.47).




a) b)

Figure 6.47: Each raters counting of plant clusters after the ACC session: a) original data, and b) log10-transformed data

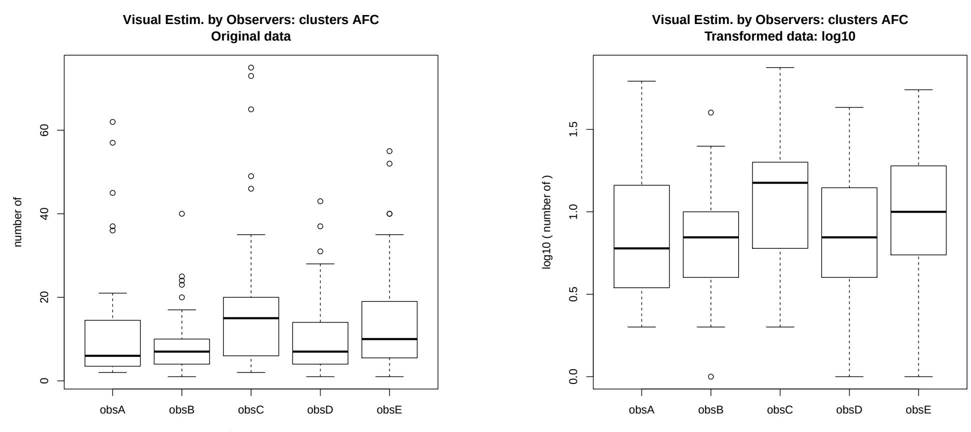
Plot-wise variation and differences are presented in figure 6.48. The maximum difference between the raters in counting the same plot was 76 plant clusters, i.e., the highest count minus the lowest count in the same plot (see figure 6.48). The median difference (over all plots) between the highest and lowest plot was: Md = 16 plant clusters and the mean: $\overline{x} = 20.65$ plant clusters.




b)

Figure 6.48: Plant cluster counts per sample plot: a) original data, and b) log10-transformed data

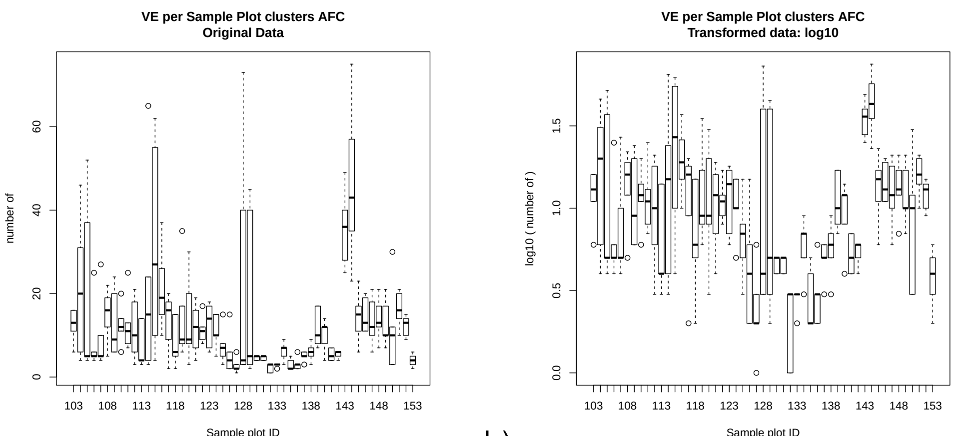
**Results of the raters' 2nd counting of plant clusters:** After the ACC-session, an ANOVA-test was conducted on the counting results. It reported a *significant* difference in count estimates in between the raters. $F = 8.093$ at $df = 4\, and\, 250$, $p = 0.000003761$. ($H_0$ was rejected at 0.05 significance level), see 6.49.




a) b)

Figure 6.49: Each raters counting of plant clusters after the AFC session: a) original data, and b) log10-transformed data

Plot-wise variation and differences are presented in figure 6.50. The maximum difference between the raters in counting the same plot was 70 plant clusters, i.e., the highest count minus the lowest count in the same plot (see figure 6.50). The median difference (over all plots) between the highest and lowest plot was: Md = 13 plant clusters and the mean: $\overline{x} = 17.8$ plant clusters.




b)

Figure 6.50: Plant cluster counts per sample plot: a) original data, and b) log10-transformed data

To assess the extent of agreement between the five raters' intraclass correlation coefficients, ICC(2,1), were calculated for the two sessions of plant cluster counting (see table 6.9.

| Count during session | Reliability as ICC(2,1) | no. of images $n$ | p-value |
|---|---|---|---|
| ACC | $\mathbf{0}.25$ | 50 | $\mathbf{3 * 10^{-7}}$ |
| AFC | $\mathbf{0.40}$ | 50 | $\mathbf{2.7 * 10^{-10}}$ |

Table 6.9: The five raters' levels of counting agreement as of the ICC(2,1)

### Intra-rater Agreements in Counting Plant Clusters

After each session of visually estimating the cover of woody plants, herbs, and grass, the raters were asked to count the number of clusters in all 51 images. As already mentioned, this procedure was carried out twice using either the ACC method, or the AFC method, so the counting occurred twice for each image per rater. Based on this counting data, an intraclass analysis was conducted to investigate whether or not each rater was in agreement with himself/herself. In other words, did the rater make the same estimate both times?

The intraclass analysis was conducted as several pairwise student's ttests (also called dependent t-tests) (Field et al., 2012, pp. 386), or matched pairs. As stated, two sets of *paired* samples were collected from each rater. The estimates were made by the same rater on the same images, firstly during the ACC session and then again during the AFC session. The null hypothesis $H_0$ was that the rater in question has a mean difference, $D$, between the paired plant cluster, counting observations that equal zero. The alternative hypothesis $H_a$ was that the mean difference is not equal to zero. For the pairwise student's t-test, the t-values on the log10-transformed data were computed using the equation at 6.8 .

$$t_{paired} = \frac{\overline{D} - \mu_D}{s_D/\sqrt{N}} \tag{6.8}$$

where $\overline{D}$ is the mean difference between the two samples, $\mu_D$ is the expected difference between the two poulation means, and $\left(s_D/\sqrt{N}\right)$ is the estimated standard error of the differences, where $s_D$ is the standard deviation of the differences in the sample and the size of the population, $N$.

$H_0 : \overline{x} = 0$ , and $H_a : \overline{x} \neq 0$

**Result:** The test results on the paired transformed log10-data show that, on average, the differences between all raters' estimates were non-significant, p > 0.05 (see table 6.10), i.e. the null hypothesis $H_0$ could not be rejected. It therefore follows that the alternative hypothesis $H_a$ could not be accepted. In other words the difference in estimates from the two plant counting sessions could not be proven (by testing) to be not significantly different.

For each table row, 50 degrees of freedom applies, $df = n - 1$ where the number of images $n = 51$. The result in differences is graphically presented in appendix H.

| rater ID | Counting after the ACC session | Counting after the AFC session | Mean of the differences, $\overline{D}$ | t(50) | p-value |
|---|---|---|---|---|---|
| A | $\overline{x} = 0.899$, $SE = 0.056$ | $\overline{x} = 0.871$, $SE = 0.055$ | 0.028 | 1.728 | 0.090 |
| B | $\overline{x} = 0.803$, $SE = 0.043$ | $\overline{x} = 0.840$, $SE = 0.044$ | $-0.037$ | $-1.7319$ | 0.089 |
| C | $\overline{x} = 1.066$, $SE = 0.055$ | $\overline{x} = 1.077$, $SE = 0.055$ | $-0.011$ | $-0.6764$ | 0.501 |
| D | $\overline{x} = 0.885$, $SE = 0.050$ | $\overline{x} = 0.882$, $SE = 0.052$ | 0.003 | 0.064 | 0.949 |
| E | $\overline{x} = 0.923$, $SE = 0.058$ | $\overline{x} = 1.002$, $SE = 0.053$ | $-0.079$ | $-0.978$ | 0.333 |

Table 6.10: The raters' two (paired) sessions of counting plant clusters, df=50

### 6.9.4 Conclusion

**Visual estimates of plant cover:** The ANOVA test results showed inconsistency when it came to estimating plant cover. Four out of six tests were found to be significant concerning differences in the mean estimates of cover. A summary of the results from sections 6.9.2 through to 6.9.2 is presented in table 6.11, were significant results are marked as bold.

| Target plants | VE method | Mean diff, $\bar{x}$ (%) | Median, $Md$, diff (%) | Max difference | $F(df1, df2)$ | p-value |
|---|---|---|---|---|---|---|
| Woody Plants | ACC | 19.6 | 14 | 79 | $F(4, 205) = 1.499$ | 0.2037 |
| | AFC | 9.0 | 6 | 35 | $F(4, 185) = 5.219$ | **0.0005** |
| Herbs | ACC | 30.0 | 26.5 | 70 | $F(4, 245) = 1.227$ | 0.2998 |
| | AFC | 13.9 | 9 | 44 | $F(4, 230) = 4.674$ | **0.0012** |
| Grass | ACC | 19.3 | 9 | 59 | $F(4, 180) = 2.666$ | **0.0340** |
| | AFC | 22 | 9 | 84 | $F(4, 160) = 5.427$ | **0.0004** |

Table 6.11: Summary of difference in between raters from section 6.9.2 to 6.9.2

It should be noted that some of the raters sometimes found it hard to differentiate between a pine and a tuft of grass, when viewed from the nadir perspective. This could explain some of the fluctuations in cover estimates.

The **raters' level of agreement** was assessed by computing the intracorrelation coefficient ICC(2,1). The degrees of freedom used for the calculations in table 6.12 are $df1 = (n-1)$ and $df2 = (o-1)(n-1)$, where $o$ is the number of raters, and$n$ is the number of estimated images.

| Target plants | VE method | Reliability ICC(2,1) | no. of images $n$ | p-value |
|---|---|---|---|---|
| Woody Plants | ACC | **0.27** | 42 | $\mathbf{1 * 10^{-6}}$ |
| | AFC | **0.38** | 38 | $\mathbf{5 * 10^{-10}}$ |
| Herbs | ACC | **0.36** | 50 | $\mathbf{1 * 10^{-11}}$ |
| | AFC | **0.41** | 47 | $\mathbf{4 * 10^{-14}}$ |
| Grass | ACC | **0.16** | 37 | **0.003** |
| | AFC | **0.13** | 33 | **0.010** |

Table 6.12: The five raters' level of agreement as of the ICC(2,1)

The ICC(2,1) results in table 6.12 shows values in between $ICC(2,1) = 0.13$ (when estimating the cover of grass using AFC) up to $ICC(2,1) = 0.41$ (when estimating the cover of herbs using AFC) in the interval of 0 to 1. If values around the arithmetic mean (in between 0.4 to 0.6) would be characterised as being moderate agreement, then the obtained results could be described as *poor agreement* up to just about fairly *moderate* agreement in between the raters.

**Counting plant clusters**: The results for the investigation of raters' counting of plant clusters showed instability in counting between the raters. However, when the same test were repeated twice, the single rater showed stability in counting.

*Inter-rater reliability*: An analysis of variance tests showed that there were significant differences between the raters' counting of plant clusters. This applied both to the counting after the ACC session, as well as that carried out after the AFC session. In addition, ICC values were calculated (see table 6.9),and the values were found to be similar to when the raters had to estimate the extent of cover (see table 6.12): $ICC(2,1) = 0.25$ and $ICC(2,1) = 0.40$ respectively. Again a *poor agreement* up to the lower boundary of a *moderate* agreement between the raters in counting plant clusters.

*Intra-rater reliability:* Differences between first, ACC, and second, AFC, VE sessions for raters A to E were compared using paired (dependent) students t-tests. The mean of differences, together with each session's standard errors were low (see table 6.10). This shows consistency in counting between the two paired sessions by the same rater. The results give an indication that the method in section 6.4 describing how to quantify the number of clusters was understandable and stable. Hence, the differences in counting did not result in a significant difference.

If one rater estimated either the cover or count plant clusters, the inter-rater reliability results show *poor* to *moderate* agreement between raters. Hence, caution should be exercised when interpreting individual raters' results.

## 6.10 5$^{th}$Study of Visual Assessments: Älvdals-banan, Oxberg

### 6.10.1 Method

The data was collected along the Alvdalsbanan in Oxberg, Sweden, as presented in section 5.5. For the analysis, two domain experts visually estimated the total plant cover (woody plants, herbs and grass) using the aerial foliage cover (AFC) and sub-plot frequency (SF) methods (see definitions in section 4.3). Because of the large presence of woody plants it was interesting to try these two methods on woody plants only. All visual estimates were made by considering a sample area of 1 x 1 metres, with or without the aid of a sub-plot grid. The results of the visual observations are presented in table 6.13. As there were five sample plots, each rater made five visual estimates (VE), as follows:

1. The first VE of the *total cover* was carried out using the *AFC* method, with no other assistance than a boundary of a square meter (marked with pins).

2. The second VE of the *total cover* was carried out using *AFC* method, with the assistance of a boundary of a square meter, a sub-plot quadrat sub-plot frame quadrat including a 10 x 10 sub-plot (each sub-plot = 10*10 cm).

3. The third VE of the *woody plants cover* was carried out using the *AFC* method, with the assistance of a boundary of a square meter, including a 10 x10 sub-plot (each sub-plot = 10*10 cm).

4. The fourth VE of *total cover* was carried out using the SF methd with the assistance of a boundary of a square meter, including a 10 x 10 sub-plot (each sub-plot = 10*10 cm).

5. The fifth VE of *woody plants cover* was carried out using the SF method, with the assistance of a boundary of a square meter, including a 10 x 10 sub-plot (each sub-plot = 10*10 cm).

| VE of total cover (no grid) | plot 200 | plot 201 | plot 202 | plot 203 | plot 204 |
|---|---|---|---|---|---|
| **Rater A** | 50 | 55 | 40 | 25 | 15 |
| **Rater B** | 45 | 20 | 15 | 15 | 10 |
| **VE of total cover, AFC (Grid aid)** | **plot 200** | **plot 201** | **plot 202** | **plot 203** | **plot 204** |
| **Rater A** | 62 | 37 | 16 | 15 | 13 |
| **Rater B** | 52 | 28 | 20 | 18 | 15 |
| **VE of woody plants, AFC (Grid aid)** | **plot 200** | **plot 201** | **plot 202** | **plot 203** | **plot 204** |
| **Rater A** | 30 | 21 | 15 | 14 | 9 |
| **Rater B** | 25 | 22 | 13 | 13 | 7 |
| **VE of total cover, SF** | **plot 200** | **plot 201** | **plot 202** | **plot 203** | **plot 204** |
| **Rater A** | 86 | 59 | 46 | 61 | 47 |
| **Rater B** | 94 | 81 | 68 | 73 | 72 |
| **VE of woody plants, SF** | **plot 200** | **plot 201** | **plot 202** | **plot 203** | **plot 204** |
| **Rater A** | 64 | 49 | 43 | 51 | 39 |
| **Rater B** | 34 | 52 | 49 | 50 | 41 |

Table 6.13: Oxberg: on-site visual cover examination

### 6.10.2 Results and Conclusions of the 5th Study

The difference in assessing cover between the raters was computed (see results in table 6.14). The reliability of the raters visual estimates were assessed by computing the ICC(2,1) and the Krippendorff's $\alpha$ coefficients (see results in table 6.15)

| Assessment method, what to assess, aid | Diff. plot 200 | Diff. plot 201 | Diff. plot 202 | Diff. plot 203 | Diff. plot 204 | Mean diff. |
|---|---|---|---|---|---|---|
| VE AFC, total cover, no aid | 5 | 35 | 25 | 10 | 5 | 16 |
| VE AFC, total cover, grid aid | 10 | 9 | 4 | 3 | 2 | 5.6 |
| VE, AFC, woody plants cover, grid aid | 5 | 1 | 2 | 1 | 2 | 2.2 |
| VE SF, total cover, grid aid | 8 | 22 | 22 | 12 | 25 | 17.8 |
| VE SF, woody plants cover, grid aid | 30 | 3 | 6 | 1 | 2 | 8.4 |

Table 6.14: Difference in cover estimates per plot

| Assessment method, what to assess, aid | Reliability ICC(2,1) | ICC(2,1) p-value | Krippendorrf's $\alpha$ |
|---|---|---|---|
| VE AFC, total cover, no aid | 0.42 | 0.094 | 0.291 |
| VE AFC, total cover, grid aid | 0.94 | 0.0037 | 0.93 |
| VE, AFC, woody plants cover, grid aid | 0.94 | 0.0012 | 0.935 |
| VE SF, total cover, grid aid | 0.46 | 0.016 | 0.283 |
| VE SF, woody plants cover, grid aid | -0.58 | 0.83 | -0.354 |

Table 6.15: Agreement between the two domain raters' VE

Two domain experts participated in this investigation. Five trials were conducted with five sample plots in each trial. Due to the fact that there were only five sample plots to assess per trial, it is impossible to draw any general conclusions. Some indications were given though. The reliability results indicate (not surprisingly) that it was easier if a grid assists the rater when conducting VE using AFC of total cover. The reliability coefficients of the VE by using AFC assessing both the total cover and woody plants cover (assisted by a grid), were *excellent*.

The reliability coefficients of the VE by using SF assessing the woody plants cover (assisted by a grid), was in between *moderate* to *poor* (see table 6.15). The VE differences for that trial are high $\overline{x} = 17.8\%$ (see table 6.14).

The reliability coefficient values of the VE by using SF assessing the woody plants cover (assisted by a grid), were all *negative*. This indicates that the raters estimates were worse than random. Often this indicates some kind of structural misunderstanding between the raters. By looking at the VE differences (see table 6.14) for that trial, it is quite convincing that sample plot no. 200 was the root of this misunderstanding, perhaps because of confusion of how to assess the extent of the bigger woody plants.

## 6.11 6$^{th}$ Study of Visual Estimates On-site in Vetlanda

Three domain experts provided visual estimates (VE) on-site in 12 out of the 179 sample areas (see section 5.6).

### 6.11.1 Method

Twelve sample areas were selected by a systematic sampling method in which the starting position (of the first sample area) was chosen at random, and every eighth sampling area was assessed accordingly. Each sample area was represented by a rectangular area comprising five ballast areas in between six sleepers.

In the current study, VE was conducted by estimating the ACC, i.e., by estimating the ground covered by the vertical projection of the outermost perimeter of the plant, also known as the convex hull (as described in section 4.3.1). A plant was defined as an individual. Alternatively, when it was practically impossible to identify individual plants, if there were tightly clumped groups of individual plants (i.e., plant clusters), they were counted as being one plant.

All raters agreed that an A4 sheet (21 x 30 cm) represented approximately 1% of a sample area. As the raters used an A4 sheet to record their observations they also had the opportunity to use the same sheet when conducting their estimates. No time limits were applied and the raters reported their estimates independently, i.e., without any interaction between each other.

### 6.11.2 Results and Conclusions of the 6th Study

An analysis of variance (ANOVA) was initially used to investigate differences between the raters' assessments. In addition to the ANOVA test, the intra-class correlation coefficient ICC(2,1) was calculated in order to assess reliability in terms of the consistency of measurements made by the raters (see section6.3). Each subject (represented in this investigation by the described sample areas on the railway embankment) was measured by each rater. The raters were assumed to be representative of a larger number of similar raters in the population, i.e., domain experts. A summary of the average of the three rater observations in the 12 randomised sample plots is given in table 6.16. Note that this particular railway section underwent vegetation management in between the two sessions.

| | Cover in June (%) | Cover in August (%) |
|---|---|---|
| Mean | 12.89 | 2.6 |
| Std. deviation | 1.55 | 1.8 |
| Max | 29 | 7 |
| Min | 4 | 0 |

Table 6.16: Raters average cover estimates

In order to compare the cover estimates made by raters A, B and C, two one-way ANOVAs were conducted, one for June and one for August. Both test results were significant ($p < 0.05$), meaning that, on neither occasion were the raters in agreement. It was not in the interests of this investigation to identify which rater differed from the others. Thus, no post-hoc tests were performed. To test the reliability of the raters' assessments, i.e., to see whether they were in agreement, the ICC(2,1) was calculated:

June: $ICC(2,1) = 0.53$ was *significant* on level $\alpha = 0.05$, $df = 11$, $p = 3.9 * 10^{-7}$.

August: $ICC(2,1) = 0,51$ was significant on level $\alpha = 0.05$, $df = 11$, $p = 3.15 * 10^{-6}$

The results of the ANOVA tests carried out for June and August 2013, respectively, showed a statistically significant difference between the three domain expert rater estimates (p<0.05). ICC2 coefficient values could be considered as showing *moderate* reliability for a single rater, i.e., how accurate a single rater would be if they made the estimates on their own. Again, the latter result, in conjunction with the ANOVA result, makes the reliability of visual estimates appear as weak. The results of this investigation support the hypothesis that humans are unreliable in their estimates of cover extent.

## 6.12 Conclusion and Discussions

The results from the investigations made (in section 6.6, 6.7, 6.8, 6.9, 6.10 and 6.11), show inconsistency among raters' estimates. When using humans for assessing the extent of vegetation cover, the results highlighted the importance of having a predetermined strict protocol of how to estimate cover. This would reduce systematic errors made by the misinterpretation of how to assess vegetation cover.

Regarding the raters' ability to estimate the amount of vegetation along a real railway embankment (on-site), or from images, the combined results of these investigations exhibited a pattern, indicating *insufficient reliability* and *relatively large central tendencies in visual estimate (VE) differences.*

Based on the results of these investigations, **an automated monitoring approach is suggested**, thus transferring the manual inspections into objective monitored inspections using machine vision.

# Part V

# Machine Vision and Machine Learning

Based on the conclusions in the previous part IV it was suggested that an automated monitoring approach was preferable in favour of manual assessment of the extent of vegetation. In this parts of the machine learning process are presented which addresses the research questions concerning how to reach viable solutions by making use of machine learning and machine vision.

In this thesis, there is a need of capturing enhancing the knowledge possessed by key experts. These key experts include inspection personnel, vegetation engineers, railway maintenance personnel, ecologists, foresters and botanists. The following definition of knowledge is used: *"Knowledge refers to the stored information or models used by a person or a machine to interpret, predict, and appropriately respond to the outside world*" Fischler and Firschein (1987).

# Chapter 7

# Classification in Machine Learning for Machine Vision

The motivation of this chapter is to give a quick summary of learning by classification for use in machine vision.

The goal of machine and computer vision is to extract useful information from images or video frames. Often this extraction process ends up with image classification where an image is classified according to its visual content. If an image can be thought of as being a dataset of extracted properties from the same then the process of classifying images will be a subset of *pattern recognition*.

Often the field of machine vision are broken down according to degree of abstraction from an image/or video frame as: low-level processing, mid-level processing and high level processing. These are often a chain of events in developing algorithms and there are no crisp boundaries between these abstraction layers, but they will serve as a starting point.

At the low-level a mapping takes place from pixels to pixels. Often used operations include edge detection, corner detection, scale-invariant feature transforms etc. The mid-level are mostly concerned with extracting descriptions from the image from the image descriptions extracted at the low level. When processing on the mid-level the mapping takes place from pixels to regions this often include segmentation operations. At the highest level the goal is to make use of the results from the mid-level to recognise and/or localise objects or regions by classification. The mapping occur from pixels and regions to abstract categories such as ”leaf”, ”cup”, ”glass”, ”gravel road”, ”asphalt road”, ”railway” etc.

## 7.1 Recognising Objects and the Classification Model

For many people it is obvious that (we) humans are capable of recognising objects under varying conditions. These objects may vary in form, colour, texture, etc. We can recognise objects from many perspectives, in many different places, and in different sizes. All these abilities that we take for granted are complex to formally represent in the attempt of transferring these abilities to a computer. For example, in the fields of image processing, computer vision and machine vision these different view points (when looking at an object) are represented as four transformations: scaling (i.e. different size of the object), translation (i.e the same object appears in different places in the scene), rotation and shearing (i.e. changing the shape of the object, a.k.a. skewing). Although, these transforms are quite easy to apply on objects having crisp lines defining their shape and/or content, it becomes very complex when trying to extract these objects from an image acquired from a real world scene. Such images often contain undesired noise. The noise could refer to: noise produced by the camera sensor because of poor camera configuration, or by poor preprocessing of the image (e.g. information loss because of changing resolution, size, or choice of image format), or on a higher abstraction level the noise could refer to that the scene contains many objects which are not of interest (a.k.a. background). It might not have been known what are the objects of interest at the time of acquiring the image. Especially, it becomes very hard and sometimes impossible to extract natural objects, such as leaves, out of natural scenes. This is because most of the natural objects are very heterogeneous concerning shape, colour, and texture. Humans can recognise objects that are partially obstructed from view, but when it comes to doing the same using machine vision it becomes a hard problem in trying to infer what is hidden based upon what is visible.

Efforts of emulating our human capabilities within the fields of artificial intelligence and computational intelligence are encapsulated in the domains of pattern recognition and machine learning, in where both (and especially machine vision) are about making inferences based on (sensor-) perceived data. For this purpose methods are used from statistics, probability, computational geometry, signal processing, and algorithm design. The terms machine learning and pattern recognition are in many contexts treated as being the same. It has been noted that pattern recognition though there does not need to be learning, i.e. a program can developed to recognize a pattern but does not have the ability to learn from it. Learning is an additional component where the system is capable to adapt to new data for improved performance in the future. There are three kinds of learning processes: Supervised learning, unsupervised learning, and reinforcement learning.

*Supervised learning* is about learning the relationship between two datasets with help of a supervisor. The observed data X and an external variable Y (the output) that are to be predicted. The observed data X are often called the training dataset, and the output Y is usually called targets or labels. Typically the classifier is trained with the training dataset and a set of desired target outputs. Then the classifier is tested with previously unseen input data (called validation- or test dataset). The performance is then based upon the classification error rate. Supervised classifiers include multilayer perceptrons (MLP) using a backpropagation algorithm, discriminant analysis, support vector machines (SVM), random forests, learning vector quantization (LVQ), radial basis functions (RBF) etc. In addition, there is also the very common and simple k-nearest neighbor (kNN) method which is a kind of semi-supervised algorithm. (For further information about supervised learning refer to Haykin (2009), Bishop (2006) and Duda et al. (2000)).

In *unsupervised learning* it is about finding similarities in the training data. The assumption is often that the clusters discovered will match reasonably well with an intuitive classification. Here there are no target outputs associated with each input; rather the unsupervised learner brings to bear prior biases as to what aspects of the structure of the input should be captured in the output. Common unsupervised classifiers are the k-Means clustering, fuzzy k-Means clustering. Self-organizing maps (SOM), Gaussian mixture models etc (For further information refer to Ghahramani (2004), Duda et al. (2000) and Bishop (2006)).

Any classification problem (dealing with digital images/video clip fames) to be solved by use of pattern recognition includes the following general phases:

1. Acquisition of images or video film of the objects of interest.
2. Preprocessing
3. Feature extraction
4. Classification

The *preprocessing* phase typically involves segmentation operations in where (optimally) the objects or regions of interest are isolated from rest of the things not of interest in an image. If the segmentation is being done manually (by use of interactive environment like GIMP, or Adobe Photoshop), then objects of interest can confidently be segmented (since we can see what is being segmented). If this process are to be done automatically it is crucial to be confident in that the algorithm which performs the segmentation should stop when the objects of interest are found. This to be able to continue to the next phase of feature extraction. Autonomous segmentation of non-trivial images is one of the most difficult tasks in machine vision. The preprocessing of images/frames from video clips could also include: enhancement, restoration, morphological- and colour processing etc. (For further information refer to Gonzalez and Woods (2007)).

The process of *feature extraction* should formally represent (i.e. quantify) the properties of the segmented objects/regions from above and extract the values out of these features. In the case of machine learning classification or regression the input variables which feeds the learning machine are crucial. From the preprocessing phase there are *input variables* $X$ which often are multidimensional in space. Each dimension $i$ of $X$ is denoted by $X_i$ and is often called a *feature*, *predictor*, *feature vector, variable* or *independent variable*. The set of all features $X_n$ which be used as an input to a classifier are denoted as the *feature space.* The feature variables $X$ are needed in both supervised and unsupervised learning. In case of supervised learning there are also *output variables* $Y$ as well. They symbolically shows or points out to the classifier (e.g a certain type of a neural network) what often called the *targets, response* or *dependent variable*s.

The purpose feature extraction is to represent information about an object which is capable of differentiating one object from another. Sometimes it is useful or even necessary to minimize such information by reducing those features which do not act as an discriminant in between objects. This is called dimensionality reduction. One such approach is the principal component analysis.

The extracted features are then to be processed by a classifier. (For further information refer to Nixon and Aguado (2008)).

In the *classification* phase the inputs i.e. the extracted features are divided into two or more classes by a learning process. The model produced by the learning process then assigns unseen inputs to one or more of these classes. In essence the recognition of objects are made in the classification process where a label is assigned to each object based on the provided representation from above, i.e. the feature extraction.

After several experiments, the Bag-of-Features approach was the selected approach for classifying images in this thesis (see chapter 10). Therefore a short review of Bag-of-features is motivated.

### 7.1.1 Classification using Bag-of-Features

The Bag-of-Features approach was the selected approach for classifying images in this thesis. Therefore a short review of Bag-of-features is motivated.

The approaches Bag-of-Words , Bag-of-Visual Words, Bag-of-Keypoints and Bag-of-Features (BoF) are sometimes used interchangeably. The concepts are all about recognizing objects, where as the Bag-of-Words were originally meant for representing text (see Joachims (1998)). Visual objects can be described (conceptualised) as a ”bag of words”, ”bag features” or ”bag of keypoints ”. The ”content” of the bag of words are independent features all describing what can be interpreted as an object, e.g. a train. In early 2000 the BoW approach was extended to include visual features instead of just words (Lazebnik et al. (2006), Csurka et al. (2004), Sivic and Zisserman (2003) and Nowak et al. (2006)).

BoF approaches are characterised by the use of an orderless collection of local image features.

**Image classification using the Bag-of-Features Approach in Machine Vision** In computer/machine vision, the bag-of-words model (BoW model) can be applied to *image classification*, by treating image features as visual words. In document classification, a bag of words is a sparse vector of occurrence counts of words; that is, a sparse histogram over the vocabulary. In computer vision, a bag of visual words is a vector of occurrence counts of a vocabulary of local image features.

The difficulties are that objects in our three dimensional world look different from different angles and different lighting conditions when mapping them into a two dimensions such as images or frames in a video sequence. The same type of object often appears different, e.g. see the locomotives in figure 7.1. This is known as intraclass variation. Sometimes it makes sense to specify the class more, for example the locomotives could be categorised by model number and/or by colour. Interclass variations are variations between different types of objects, e.g. locomotives and cars.

Figure 7.1: A fictive class of locomotives

Natural objects (opposed to man made objects) often have a asymmetric shape and colour and experience a high degree of intraclass as well as interclass variations. In order to let a machine learn the features of for example a locomotive a large set of training images containing locomotives should be presented. In addition in the BoF approach a set of so called negative images should also be presented. In the case of locomotives, this negative category should not contain any locomotives. As always when it comes to images/video frames the problem of separating the foreground, i.e. the object(s) of interest, from the background is often present. The BoF approach deals with image classification as follows. The BoF is generally made up of a *pipeline* that includes the following steps:

1. Feature extraction, including feature detection (feature sampling) and description

    - First, features are to be detected. These features are often corners. This *feature detection* is performed as a sampling process from the training or test image. Local sub images around the detected features are then extracted. Each part becomes a feature of the original image and are then represented using for example: Scale Invariant Feature Transform (SIFT) Lowe (2004), Dense SIFT (DSIFT), or Multi-scale Dense SIFT (MD-SIFT) Bosch et al. (2007), as used in this thesis (see 10). The result is denoted as feature vectors or feature descriptors.

2. Codebook formation and image representation

    - The features are clustered into k number of clusters using k-means clustering. Optimally, similar features are clustered together with other similar features. The similarity is dependent on the quality of the representation, as in step 1 above. The clustered features in this step build up the *visual vocabulary*, which is the content of the *codebook*.
    - The image representation part involves the mapping (i.e., assignment) of each feature to the most representative visual word (in the visual vocabulary) This process is known as vector quantisation (VQ). Alternatives to VQ include sparse coding (SC) (J. Yang et al., 2009) and *Locality-constrained Linear Coding* (LLC) proposed by (Wang et al., 2010).

3. Learning and recognition

    - The use of discriminative methods, or classifiers such as. support vector machines (SVM) with linear or non-linear kernels, or k-NN classifiers to learn category models, or classifiers from a training set.

## 7.2 Conclusion

This chapter has summarised a thorough review of machine learning and the BoF development and history. For further information on BoF development refer to O'Hara and Draper (2011) and Tsai (2012) and for machine learning refer to Bishop (2006) , Duda et al. (2000) and Prince (2012).

# Chapter 8

# Quantification of Vegetation using Machine Vision

This chapter describes various approaches used to quantify the vegetation that grows along railway embankments. The described investigations often refer back to the investigations made for human visual estimates presented in section 6. The methods used for collecting the data used in the investigations in this section can be found in chapter 5.

In this investigation, the plant attribute cover (see section 4.3.1) has been used because of its common usage, and because of the relative ease of transferring its concepts into images. Of equal importance, cover has a linear relationship, which reflects the actual amount of aboveground biomass for low, open herbaceous plants growing in low nutrient and low moisture soil (Rottgermann et al., 2000). It is assumed that those conditions are often similar to the environment found on a railway embankment.

## 8.1 Quantification of Vegetation using Machine Vision

This investigation is associated with the investigation of human visual estimates presented in section 6.11. Data collection from Vetlanda, Sweden is described in section 5.6.

It should be noted that the algorithm described in this section is not dependent on the data set from Vetlanda (Sweden). Indeed, any image data set can be used, so long as it follows the guidelines given in this section.

### 8.1.1 Algorithm Description

The only manual steps needed here are to unload images from the image sensor (i.e., the camera), and select which of them to analyse. The images should have been taken (optimally) from a nadir perspective(see figure 8.1). Thereafter the selected sequence of images will automatically be analysed as described in the steps below.

Figure 8.1: View over the trackbed

Initially, the algorithm will determine an area of interest in the red, green and blue-coloured (RGB) image in which the measurements will take place. It is important that both of the rails can be identified in the image. If so, then a fixed dimension is known, namely the nominal standard track gauge (e.g. 1435 mm between the inside of the rails). By knowing this, one can calculate an estimate of the ratio of pixels per meter in reality.

The algorithm will detect if the camera has been configured incorrectly. This happens when it cannot detect both of the rails in the image being processed.

In this thesis, the relevant track images were initially resized to 339 × 500 pixels using bilinear interpolation so as to reduce the computational burden for any further analysis.

The original image (as shown in figure 8.1), was then segmented using mean-shift clustering (Comaniciu and Meer, 2002); (Fukunaga and

Hostetler, 1975). This method segments the image into a number of coloured clusters, as in figure 8.2a.

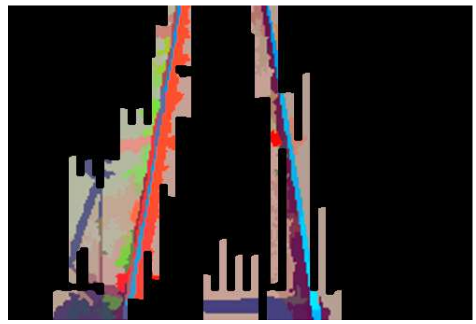
a)  b) 

Figure 8.2: a) After Mean-Shift Clustering, and b) HOG mask on top of the decorrelated stretch result

When a mean-shift algorithm is used to segment images essentially three parameters have to be set: 1) the spatial resolution, *hs*, which affects smoothing and is chosen according to the size of the image and objects; 2) the range resolution, *hr*, which affects the number of segments/clusters. If the image contrast is low, then the *hr* value should be kept low as well; and 3) the size of smallest segment, *M*. This value should be chosen based on the size of noisy patches. Experimentally, these three parameter values were set to: *hs* = 32 and *hr* = 4 and *M* = 30.

At this stage, both of the rails are clearly visible, but they are not yet represented as objects in any way (e.g. numerically represented as colour, coordinates, angles, etc.). A decorrelation stretch was further applied to enhance the difference in colours in the RGB image.

In parallel with the mean-shift clustering, another segmentation procedure, known as the *Histogram of Oriented Gradients (HOG)* (Dalal and Triggs, 2005), was also carried out to guarantee the detection of rails.

HOG is typically used to detect objects from images using segmentation. The motivation for this operation is to create an image mask that hides everything but the rails. It should be noted that HOG segmentation was carried out on the original RGB image. The HOG segmentation output was further processed by a number of morphological operations (see, for example, a description of such operations Gonzalez and Woods (2007)).

These yielded a resulted in a binary image (black and white). The binary image, called the HOG mask, was then placed as a layer on top of the resulting image from the mean-shift segmentation in order to be able to automatically segment out the rails in the image (see figure 8.2b). Identification of the rails in the image is deemed to be important, because it allows the nominal track gauge to be calculated (typically 1435 mm between the inside of the rails). By knowing this, one can calculate the ratio of pixels in the image and map it to reality.

It is important that both of the rails can be identified in the image. The green and blue colour channels contain most colours that represent steel. Thus, the red channel was disregarded, and only the green and blue channels were kept. For further image enhancement, additional greyscale morphological operations were implemented. Finally, in this step, the greyscale image was converted into a binary image by Otsu thresholding (Otsu, 1979). This dynamically calculates a threshold value for which grey level pixels should be white and which black (see figure 8.3a.

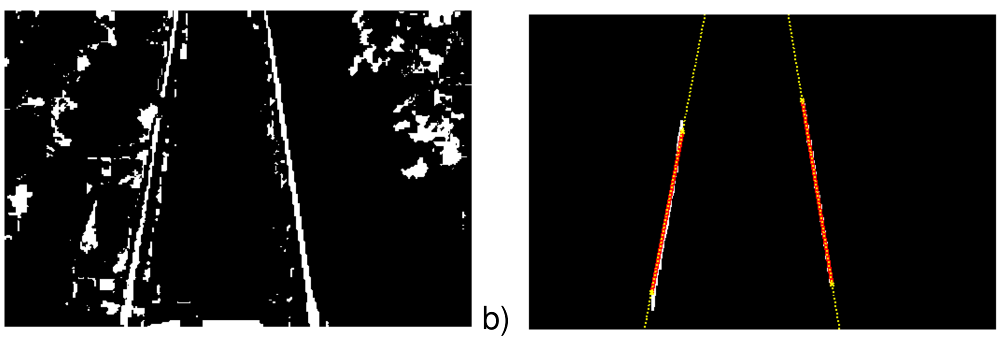
a)  b)

Figure 8.3: a) After Otsu thresholding, and b) Extrapolated skeletonised line segments representing the rails

Next, all the white objects were skeletonised using morphological skeletonisation. By using this operation, all areas will be represented as linear structures. Again, in order to hide everything but the rails the HOG mask was placed as a layer on top of the skeletonised image. Only the longest line segments in the HOG mask were kept. These lines were considered as rail candidates and were extrapolated, as seen in figure 8.3b.

As the rails were identified, a mask was created, representing a region of interest (ROI). This mask was based on the line segments (i.e., the rails) after skeletonisation. The ROI is the area in which vegetation will be sampled and later measured. An example of the ROI mask placed on the original image can be seen in figure 8.4.

The size and relative position of the ROI is based upon the fixed positions of the two rails. These two attributes can be altered by, for example, a national railway administration, depending on which area (of the embankment) they want to measure. As an example, the national railway administration in question could regulate that the common sampling area should be the area inside the rails plus the area two meters either side of the rails. Subcontractors will then conform to this regulation, so that everyone measures the same objective area. The algorithm detects the rails; thus, any arbitrary area in an image can be set as a sampling area.

The perspective projection in images means that all items appear smaller when they appear at the top of the image rather than at the bottom of the image. Thus, a perspective correction was implemented, as shown in figure 8.5a.

Figure 8.4: ROI mask layer placed on top of original image

As a result of this step, the rails will appear parallel. Any object in the upper parts will be distorted when they are stretched out, and will therefore also become blurry. However, all objects will approximately be in the same scale.

The amount of vegetation was computed from these perspective corrected images. Vegetation was segmented mainly via colour segmentation in the HSV colour space (Hue, Saturation, and Value), see figure 8.5a. The extracted vegetation was represented in white in a binary image and the plant cover (%) was obtained by using equation 8.1:

$$Cover = \frac{(no.\ of\ white\ pixels)}{(no.\ of\ white\ pixels) + (no.\ of\ black\ pixels)} * 100 \qquad (8.1)$$

a) b)

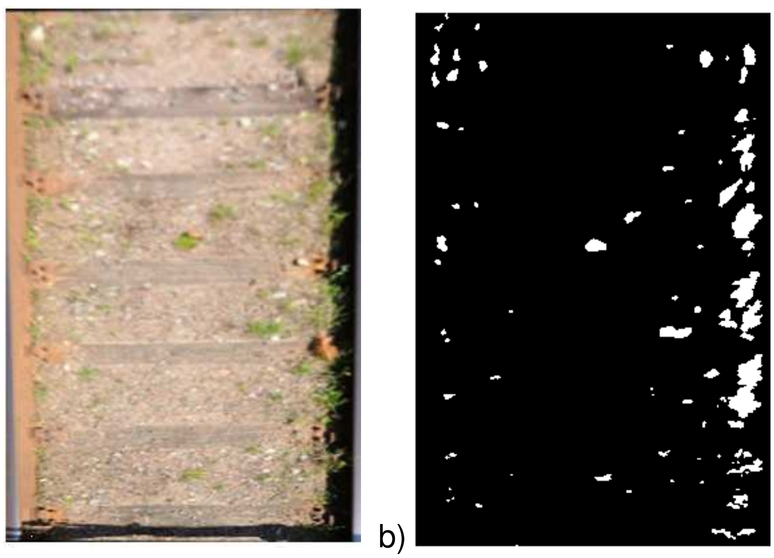


Figure 8.5: a) Perspective corrected image, and b) Extracted vegetation

### 8.1.2 Machine Vision Algorithm Results

The machine vision algorithm was capable of processing 98% of the total set of images (171 out of 176 images from the session in June, and 176 out of 178 images from the session in August). The failure to process the remaining 2% of cases was attributed to the algorithm's inability to find the rails within the image.

The machine vision algorithms computed terrestrial plant cover over the entire railway section using the documented images as input, as reported in table 8.1. The values in August are significantly lower because of a very dry summer. In addition vegetation control actions were conducted in parallel with this study.

| | Cover June (%) | Cover August (%) |
|---|---|---|
| Mean | 3.19 | 0.51 |
| Std. deviation | 1.55 | 0.55 |
| Max | 13.3 | 3.76 |
| Min | 0.63 | 0 |

Table 8.1: Computed cover (%) by the algorithm along the entire railway section

### 8.1.3 Comparison: MV Algorithm vs. Manual Visual Estimates

In this section, the result of the machine vision algorithm is compared with the results of the investigation of visual estimates given in section 6.11.

The average of the three domain expert observations in the 12 randomised sample plots is summarised in table 6.16. This will be compared with the MV estimates in table 8.1. The differences in means of all plots between the human VE's and the cover computed by the MV algorithm are summarised as: $diff = \overline{x}_{humanVE} - \overline{x}_{MV}$ equalling $12.89 - 3.19 = 9.7\%$ mean difference per plot for the session in June, and $2.6 - 0.51 = 2.09\%$ mean difference per plot for the session in August.

The estimated cover reported by the three raters in the 12 randomised sample plots can be seen in figure 8.6a and 8.6b, as well as the corresponding estimates of the MV algorithm. The black dots represent the computed cover by the MV-algorithm for the same sample plots (denoted as cvSys in the figure plot legends).

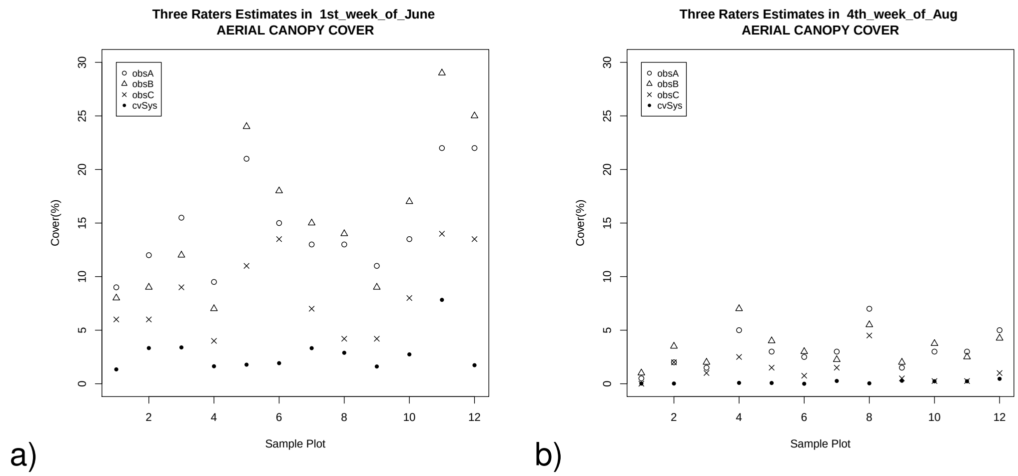
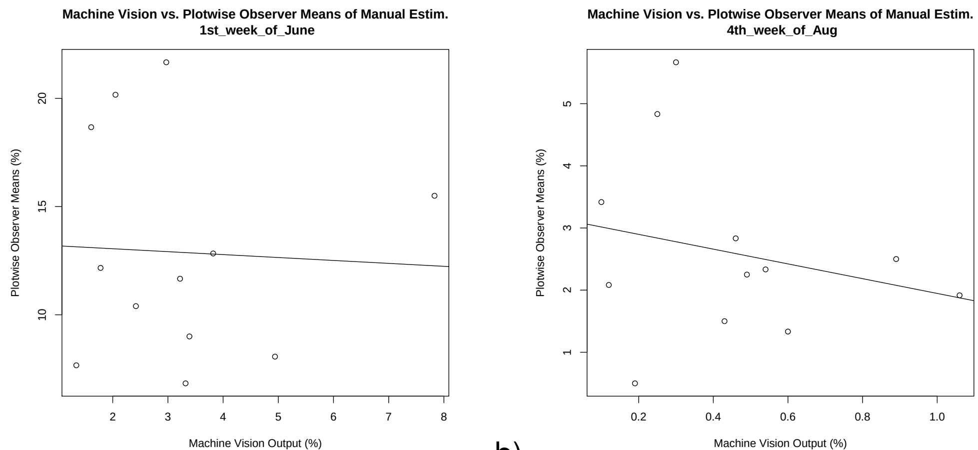



Figure 8.6: Cover estimates in a) June, and b) August

The degree of correlation between the mean for all raters plotwise compared with the machine vision computations for the same plots was calculated (see figure 8.7a for June, and 8.7b for August).

The Pearson product-moment correlation coefficients were computed to show the degree of association between the two variables: the mean for raters plotwise and the MV computation, respectively.

The coefficients showed that there was a non-significant random correlation between the mean for the raters plotwise and the MV computation for the same plots, $r = -0.05$, $df = 10$, $p = 0.883$ for June (where df is the number of data points minus 2). For August, the computed correlation showed on a non-significant very weak negative correlation, $r = -0.24$, $df = 10$, $p = 0.45$. The non-significance indicate that the null hypothesis ($H_0$) cannot be rejected, where $H_0$: *The true correlation is* $r_0 = 0$, i.e. random. The alternative hypothesis$H_a$: The actual correlation of the population is $r_a$, which is not equal to $r_0$.




a) b)

Figure 8.7: Correlation manual VE vs. MV in: a) June, and b) August

### 8.1.4 Conclusion

This investigation has presented a machine vision approach to automating the process of detecting vegetation on railway tracks. For this purpose, 179 track images were acquired. The results achieved in the current work have shown that the use of image data for detecting vegetation is indeed possible and that such results could form the base for decisions regarding vegetation control. It is also worth mentioning that the results were evaluated by comparing them with human visual estimates for the sake of validation. An objective measurement such as the one proposed in thesis not only offers easy access to the measurements to all the involved parties, but also makes the subcontracting process easier i.e., both the subcontractors and the national railway administration are given the same reference framework concerning vegetation before signing a contract, which can then be crosschecked post maintenance. Biological diversity along the embankments can be mapped, and maintained through better, and robust monitoring procedures.

The results show that the human visual estimates were not in agreement with each other; indeed, they often exaggerated the extent of vegetation cover when compared with the machine vision's output. These results are strengthened by the relatively weak ICC2 coefficient results at ICC(2, 1) =

0.53 and ICC(2, 1) = 0.51.

Upon comparing the raters' plotwise mean estimates with the machine vision output, the results show that the human visual estimates do not correlate with the results reported by the machine vision output. That is to say, there was a low degree of association between the human estimates and the MV estimates. As such, the results indicate that it is difficult to fit human estimates by regression with the machine vision result. The results in this investigation were non-significant. However, other studies (Klimes, 2003); (Benavides and Jesús, 2009); (Nyberg et al., 2013b); and (Nyberg et al., 2013a) have also identified problems with human visual estimates, all indicating that one should take care when interpreting visual estimates assessed by humans. Any one person (e.g. a railway inspector) can be consistent in his/her judgements, but when comparisons are made with other people, it may lead to misleading results.

# Chapter 9

# Estimating the Quantity of Biomass along Railway Sections

**Motivation**: Measuring biomass is central in many investigations (see section 4.3.4). Sampling biomass is very expensive and therefore in the long run often inefficient. The biomass compromises the root and the shoot of a plant. In terms of biomass, it has been shown that there are correlations between roots and shoots (see section 3.4).

The purpose of this investigation was to investigate if there is a correlation between the dry weight of the root and the dry weight of the shoot in the *special environment that exists on railway trackbeds/embankments*. If a correlation does exist, then root weights can be estimated by just by sampling shoots. Now, sampling shoots is also time consuming and ineffective, so the next step is to investigate if there is a correlation between the shoot and the aerial plant cover. This has been shown earlier (see 5.2), but can it be confirmed on the trackbeds?

## 9.1 Correlation between Root and Shoot Weights and Plant Cover

**Method**: The data used in this investigation comes from two separate data collection occasions and different field layer classes.

The first data set (hereafter referred to as the *Grycksbo data set*) comprised ten sample plots (plots 1 to 10) on the Falun-Grycksbo railway (Sweden) and is described in detail in section 5.3. This data foremost comes from the *Bilberry and Grass* field layer classes and are described in table A.1 (see details about dominant field layers in section 5.2). Both roots and shoots were harvested.

The second data set (hereafter referred to as the *Oxberg data set*) were collected on the railway that passes through Oxberg (the so-called Älvdalsbanan, in between Mora and Märbäck, Sweden) which is about 100 km further North-West from the first data collection point (sample plots 1 to 10 on the Falun-Grycksbo railway). Details about the data collection can be found in section 5.5. In Oxberg, *Lichens* can be characterised as the dominant field layer type. Conifer trees are the dominant woody plants in the surrounding forests, as well as on the railway embankment (see figure 9.1).

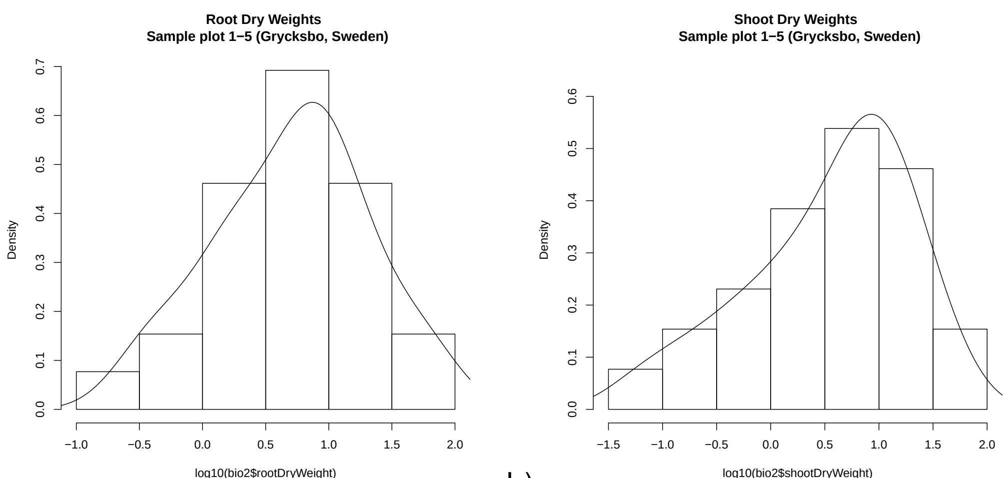
Figure 9.1: Overview, plot 200-204 (Oxberg, Sweden)

For the analysis, the data sets were log10 transformed by taking the log10 of each observation, i.e., the root and shoot data, respectively. After the log10 transformation, the data were normally distributed. The goal is to fit, by regression, the allometric formula (as presented in equation 9.1).

$$log_{10}(rootDryWeight) = b + \alpha * log_{10}(shootDryWeight) \qquad (9.1)$$

, where the coefficient $\alpha$ is the slope of the line and $b$ is the intercept of the line on the y-axis.

The transformed data from Grycksbo for sample plots 1 to 5 (weights of both roots and shoots) are presented in figure 9.2.




a) b)

Figure 9.2: Histogram and density plots of a) root dry weights and b) shoot dry weights from sample plots 1 to 5 (Grycksbo, Sweden)

The transformed data from Oxberg are presented in figure 9.3 for sample plot with 200 to 204 (both root and shoot weights).

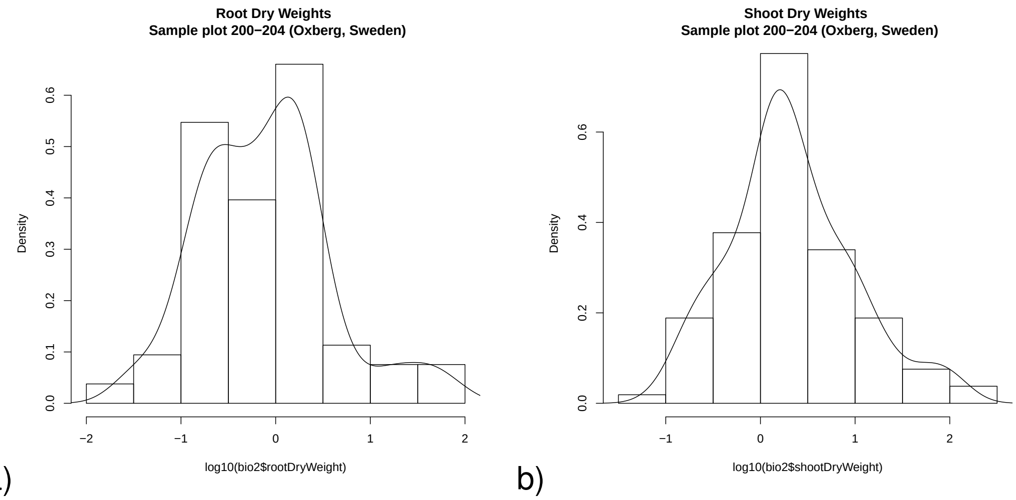



Figure 9.3: Histogram and density plots of a) root dry weights, and b) shoot dry weights from sample plots 200 to 204 (Oxberg, Sweden)

## 9.2 Results

In order to estimate, for example, the root dry weights by knowing the shoot dry weights (and, by so doing, avoiding the need to tedious work of excavating the roots), a regression was performed on the ***Grycksbo data set*** for the *Bilberry and Grass field layer class*.

As seen in figure 9.4, the shoot dry weight was a significant predictor of the root dry weight. The regression analysis for the log10 transformed the Grycksbo data set, resulting in the following estimates of $b$ and $\alpha$ (see equation 9.2).

$$log_{10}(rootDryWeight) = 0.72360 + 0.31399 * log_{10}(shootDryWeight) \tag{9.2}$$

$R^2 = 0.7769,\ F(1, 24df) = 83.57,\ p = 2.753^{-9}$

The Pearson product-moment correlation between root dry weights and shoot dry weights is given by the square root of the $R^2$ value (i.e the coefficient of determination): $r = \sqrt{R^2} = 0.88$ indicating a high degree positive of correlation.

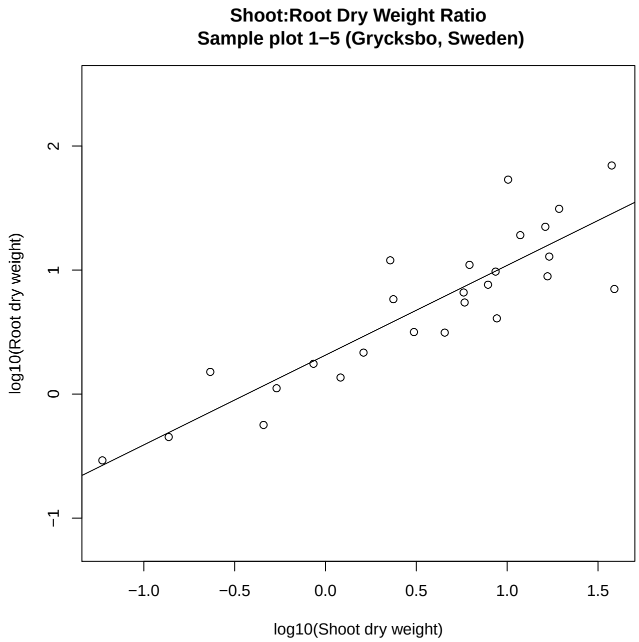



Figure 9.4: Root dry weight vs. shoot dry weight data (Grycksbo, Sweden)

For the ***Oxberg data set***, belonging to the field layer class *Lichens*, a regression analysis was conducted to predict the root dry weight from the shoot dry weight.

This time, the shoot dry weight was a *significant* predictor variable of the root dry weight (see figure 9.5 ). The regression analysis for the log10 transformed the Oxberg data set, resulting in the following estimates of $b$ and $\alpha$, see equation 9.3.

$$log_{10}(rootDryWeight) = -0.40790 + 0.90218 * log_{10}(shootDryWeight) \tag{9.3}$$

$R^2 = 0.7365$, $F(1, 104df) = 290.7$, $p = 2.2^{-16}$and $r = \sqrt{R^2} = 0.86$ indicating a high positive degree of correlation.

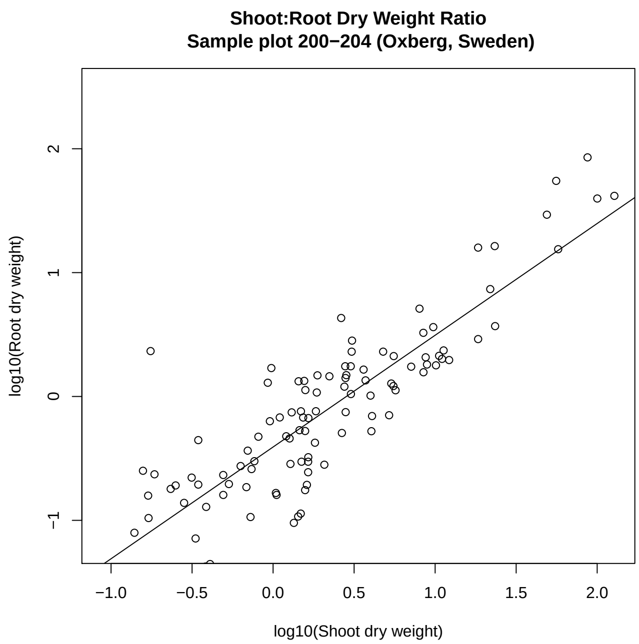



Figure 9.5: Root dry weight vs. shoot dry weight data (Oxberg, Sweden)

**Correlation between total shoot dry weight and machine vision estimate of cover in images** For the ***Oxberg data set***, belonging to the field layer class *Lichens*, a regression analysis was conducted to predict the total shoot dry weight per sample plot from machine vision estimate of cover. In all 197 plants were harvested from the five plots.

The machine vision estimates were obtained by using the algorithm described in chapter 8. The computed results is presented in table 9.1.

| sample plot ID | Cover estimate by MV(%) |
|---|---|
| 200 | 29.3 |
| 201 | 22.1 |
| 202 | 19.4 |
| 203 | 23.7 |
| 204 | 11.5 |

Table 9.1: Computed cover (%) in sample plot 200 to 204 (Oxberg, Sweden)

The total shoot dry weight per sample plot was a *significant* predictor variable of the MV computed cover from images of the same plots (see figure 9.5 ). The regression analysis for the log10 transformed Oxberg data set and for the log10 transformed cover estimates, resulted in the following estimates of $b$ and $\alpha$, see equation 9.4.

$$log_{10}(coverEstimByMV) = 0.71349+1.41144*log_{10}(totalShootDryWeight) \tag{9.4}$$

$R^2 = 0.8711$, $F(1, 3df) = 20.28$, $p = 0.02045$ and $r = \sqrt{R^2} = 0.93$ indicating a high positive degree of correlation.

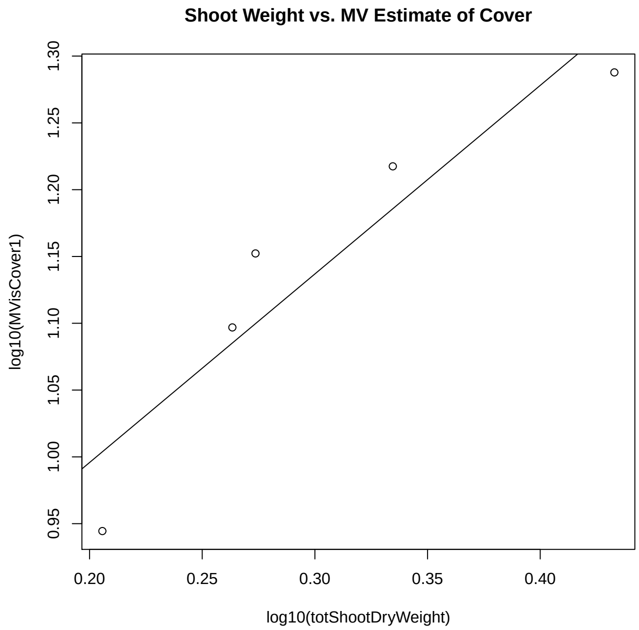



Figure 9.6: Total shoot dry weight (from 197 plants) vs. MV computed cover for the plots 200 to 204 (Oxberg, Sweden)

## 9.3 Conclusion

For the purpose of estimating the vegetation extent, which serves as a base for maintenance decisions, the results are interesting.

The collected raw data in this investigation were the dry weight of plant clusters and individual plants that can be found in areas where the Bilberry and Lichens field layer classes are dominant. Here, these classes corresponded to the Grycksbo and Oxberg data sets, respectively. The collected data were on a high aggregated level. It is important not to step down from this level if we are to draw conclusions about, for example, individual plants, or a specific species. These conclusions may be correct, but are only weakly supported by the aggregated data. This erroneous phenomenon is known as the *ecological fallacy*. Essentially, this means that one (erroneously) assumes that observed relationships for groups can be reduced and also assumed to be valid for individuals Freedman (2001).

Based on the collected biomass data, the results show that it is possible to use regression to make inferences about the root dry weights by knowing the shoot dry weights. It should be noted that a more extensive study is required to draw more general conclusions, based on the same premises, i.e., to draw conclusions based on collected plant clusters and individual plants, and any type of vegetation.

The results also indicate that it is possible to use regression to make inferences about the *total shoot dry weight* in a sample plot by computing the total cover using the algorithm in chapter 8. More extensive studies are required to draw more general conclusions.

Although it took almost two weeks to collect and prepare the vegetation data for analysis, the amount of data were of a relatively small scale. More data needs to be collected. It would be of interest to extend the investigation further by adding more field layer classes. Furthermore, it would be interesting to investigate correlations between the estimated cover extent from snapshot image data (or frames from video-clips) and plant biomass.

# Chapter 10

# Recognition and Characterisation of Woody Plants using Machine Vision

The *primary goal* of this work is to gain knowledge about the presence of large quantities of woody plants; thereby enabling the maintenance management in deciding whether to send out personnel to mechanically control the plants in question. Legal herbicides (e.g. Glyphosate in Sweden) will not kill woody plants; they are particularly ineffective when used on conifer plants (see section 1.3). In any case, the plants that die in the process contribute through the addition of a certain amount of biomass to embankments, which eventually turns into nutritious soil. A *secondary goal* is to inform the maintenance management that woody plants of certain extent and type are growing at a particular spatial location.

The objective of this investigation was to recognise and classify a data set of conifer and deciduous plants. For this purpose several classification methods were involved in the experimental phase and the Bag-of-Features method was found to present the most stable classification performance (see 7).

## 10.1 Method

In this investigation, the data collected during the indoor laboratory experiments were used (see section 5.8). In the training set there were both positive training images containing an object class (i.e. leaves or conifer seedling), and negative images which did not contain any objects of interest (see table 10.1). Training and testing procedures depend on the object class that has the least number of images. Therefore the ratio of training and test images was computed dynamically by first checking how many images there were in each object class and then choosing the class having the fewest. After this 65% percent (rounded downward to its nearest integer) were randomly chosen to belong to the training image set and the rest to the test image set. In this particular investigation (based on the number of images presented in table 10.1) this rendered 59 images as the lowest number (since Picea abies has the fewest number of images). Out of 59 images, 38 were assigned as training images and 21 were assigned as test images. These number of training and testing images were then randomly chosen (without replacement) for every object class.

| Latin name | English name | Number of images |
|---|---|---|
| Picea abies | Norway Spruce | 59 |
| Pinus sylvestris | Scots Pine | 61 |
| Betula pubescens | Downy Birch | 70 |
| Betula pendula | Silver Birch | 60 |
| Alnus incana | Grey Alder | 60 |
| Alnus glutinosa | Alder | 60 |
| Populus tremula | Aspen | 67 |
| Quercus robur | Pedunculate Oak | 70 |
| Salix caprea | Goat Willow | 65 |
| Prunus padus | Bird Cherry | 80 |
| Acer platanoides | Norway Maple | 60 |
| Negative images | - | 168 |
| *Total* | | *880* |

Table 10.1: The categories to be classified

## 10.2 Results

In this particular case, the BoF approach incorporates the following classifiers: nearest neighbour classification (1-NN) using L2 distance (NN L2), support vector machine with a linear kernel (SVM Linear), support vector machine with a linear kernel using a coding scheme called Locality-constrained Linear Coding (LLC) (Wang et al., 2010) instead of the common vector quantisation (VQ) (SVM Linear LLC max-pooling), and support vector machine (SVM Chi2).

For the k-means algorithm, the number of clusters was experimentally set to $k = 500$, where $k$ is the number of code words. The feature extraction of the training and test image set were performed by SIFT, DSIFT, and MDSIFT, respectively, as shown in table 10.2. Examples of detected image features are visualised in figure 10.1 a to d, where 10.1a is a Norway Spruce, 10.1b is a Silver Birch leaf, 10.1c is a negative image and 10.1d is an Alder leaf.

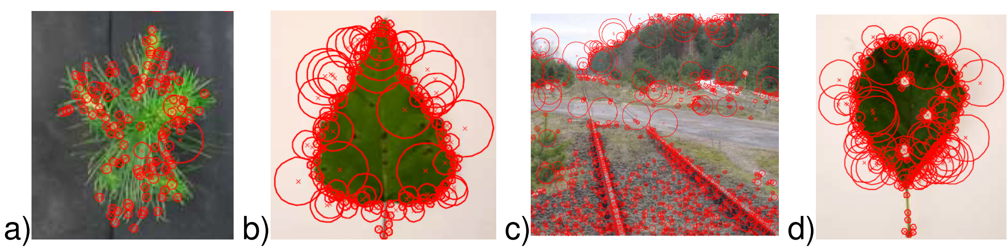



Figure 10.1: Detected image features on: a) Norway Spruce, b) Silver Birch leaf, c) negative image, and d) Alder leaf.

| | Accuracy (%) | Feature extraction method | Avg. Execution Time (sec) |
|---|---|---|---|
| NN L2 | 74.5 | SIFT | 18 |
| SVM Linear | 85 | SIFT | 21 |
| SVM Linear LLC max-pooling | 85.5 | SIFT | 70 |
| SVM Chi2 kernel | 88 | SIFT | 24 |
| NN L2 | 88.5 | DSIFT | 80 |
| SVM Linear | 88.4 | DSIFT | 91 |
| SVM Linear LLC max-pooling | 94.5 | DSIFT | 285 |
| SVM Chi2 kernel | 95 | DSIFT | 87 |
| NN L2 | 91.5 | MSDSIFT | 208 |
| SVM Linear | 84.3 | MSDSIFT | 224 |
| SVM Linear LLC max-pooling | 93.7 | MSDSIFT | 950 |
| SVM Chi2 kernel | 95.5 | MSDSIFT | 203 |

Table 10.2: BoF classification results

The confusion matrix for the alternative with the best accuracy (the SVM Chi$^2$kernel with DSIFT feature extraction) is presented in table 10.3. The matrix values represent the classification accuracy percentage for each class, A to L. The columns from left to right, denoted A to L, as well as the rows A to L are A: Alnus incana, B: Acer platanoides, C: Alnus glutinosa, D: Betula pendula, E: Betula pubescens, F: Negative image, G: Picea abies, H: Pinus sylvestris, I: Populus tremula, J: Prunus Padus, K: Quercus robur, and L: Salix caprea. In table 10.3, the columns denote the predicted species by the classifier. The rows denote the true species, i.e. the true class, as seen and assigned by a human.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 0 | 0 | 0 | 81 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 5 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| F | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| G | 0 | 0 | 0 | 0 | 0 | 0 | 91 | 9 | 0 | 0 | 0 | 0 |
| H | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 91 | 0 | 0 | 0 | 0 |
| I | 0 | 0 | 0 | 5 | 9 | 0 | 0 | 0 | 86 | 0 | 0 | 0 |
| J | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 5 |
| K | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 95 |

Table 10.3: Confusion matrix (%) for the SVM Chi$^2$kernel with DSIFT feature extraction.

## 10.3 Conclusion

Bearing in mind that natural objects are often more heterogeneous within their own class rather than outside it the results present a stable classification performance. All classifiers (except for the NN L2 with the SIFT) recorded a performance of over 84%. The SIFT feature extraction was faster than the other two approaches in particular the MDSIFT. Based on the results the MSDSIFT feature extraction together with the SVM Chi$^2$ kernel classifier exhibited the best performance. One drawback was the long execution time, averaging $\overline{x}$ =203 seconds out of ten runs. If time constraints apply, then it might be better to use the DSIFT feature extraction together with the SVM Chi$^2$kernel classifier, which only used 39% of the time, $\overline{x}$ =80 seconds.

# Part VI

# Conclusions and Discussion

# Chapter 11

# Conclusions and Discussion

This thesis has investigated the problem posed by vegetation growing along railway tracks and how it is dealt with by railway administrators. The strategy for solving the problem was presented (see section 1.3). A short review of maintenance fundamentals and how terrestrial vegetation is being measured by experts (like biologists) were investigated (see chapter 3 and 4).

It was described how the data collection was carried out (see part III). After that, investigations of how reliable humans are at estimating the extent of vegetation were described (see part IV). Next, it was described how machine vision and machine learning, as well as statistical inference could be applied for quantifying vegetation based on image data (see part V). And now finally it is time to summarise it all in this chapter.

In the beginning of this thesis, a number of research questions (RQ) were stated (see section 1.4). The answer to those questions were to targeted solutions or a deeper understanding of the problems of having vegetation growing on the trackbeds or embankments. Based on the already presented results, including the cited literature of previous work, the conclusions are now presented together with a recap of the RQs.

***RQ 1 How are railway inspections carried out with regard to the assessment of the extent of vegetation, and what methods are used for measuring vegetation along railways?*** In Sweden there are two types of inspections, namely: *safety inspections* and *maintenance inspections* (see section 1.2). Inspecting vegetation is a part of both types.

The inspections (which should form the base for maintenance decisions) are carried out manually, or none at all. Typically, inspectors are employed by subcontractors. These inspectors walk along the tracks and judge for themselves the extent of the vegetation and its condition. At their hand there are a few manuals provided by the STA.

Typically there are no objective methods for the inspection of vegetation. Hence, there is no measuring of the quantity of vegetation. The vegetation inspections carried out are subjective; the inspectors are left to exercise their own good judgement. The inspectors have little support from the administrative manuals.

For trackbeds/embankments the strategy of periodic maintenance is applied. This takes place at periodic intervals, most often by use of specially equipped (herbicide) spray trains.

Condition based maintenance (CBM) (including condition monitoring) is an important area to address (see chapter 2); bearing in mind the economic viability and the environment. Decisions for vegetation management should be based on the analysis of the data collected during condition monitoring. This could extend the life expectancy of railway objects, such as trackbeds/embankments.

**RQ 2 How is the extent of vegetation measured?** Plant species can be described by a number of quantitative features called *vegetation attributes*. Such attributes describes how much, how many, or what kind of plant species are present.

In general, the most commonly used attributes when monitoring are: *cover*, *density*, *frequency* and *biomass* (see chapter 4). All attributes are more or less applicable for measuring vegetation in the railway domain. The measure biomass is very costly to acquire directly in terms of effort, time and consequently money. The formal definition of density attribute requires individuals to be recognised and then counted. The recognition of first species and then individual could present problems even for the expert, especially when the vegetation in question is dense. Foremost cover and frequency are applicable attributes. The most common approaches for manually measuring cover are to use visual estimates (VEs) in plots, line interception, and point interception. Measuring the cover attribute can be transformed into into a problem in machine vision.

The sub-plot frequency method was found to be appealing. This is because it is easily mapped to cover; and cover in its turn is correlated to the amount of biomass. When using the sub-plot frequency method it is possible to determine presence or absence in each sub-plot and then count the number of sub-plots having the target vegetation present, even for a layperson or a computer interpreting images from a nadir view.

Plants in any geographical area are spread out in one of these three spatial patterns: random, aggregated (clumped), or uniform. There is a need to know the current spatial pattern in order to be able to make proper inferences and estimate the the quantity or amount of vegetation. For example systematic sampling along the tracks these patterns may change as the sampling process progresses forward along the tracks (see figure 11.1), and therefore it is important to evaluate the spatial pattern at each sampling point (sample plot); this to be able to make an estimate of the quantity of vegetation in the section which is between the applicable sample plot and the next sample plot (i.e. the green areas between the sample plots in figure 11.1) in this iterative process.

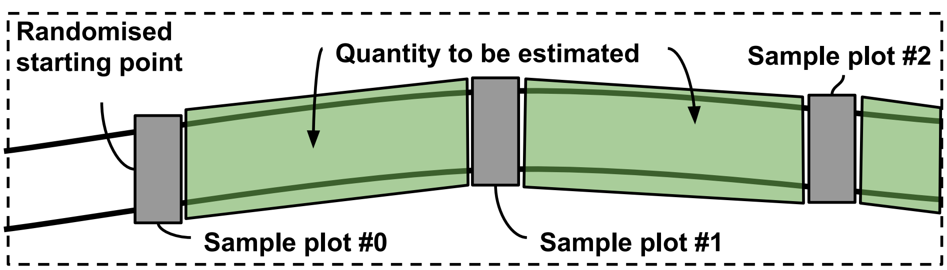



Figure 11.1: Systematic sampling along the tracks

For the transfer to machine vision, this way of measuring plant cover and frequency seems to be most useful. In addition a raster image is a grid in itself, where each square at the lowest level is a pixel, the sub-plot frequencies were found to be both interesting and useful. This is because a raster image in itself is a very fine granulated grid, depending on the resolution. For example an image size of 800 x 600 pixels comprises 480 000 potential sub-plots.

A problem with the density attribute is that it requires individual plants to be recognised. By redefining the protocol for what to count when measuring density, this attribute could be useful. For example, if a definition of plant clusters can be obtained, then a modified density attribute could involve both clusters and individuals.

With regard to production attributes, measuring plant heights was found to be highly applicable. This could be carried out using optical laser measurements from above.

To be able to describe and predict plant populations on railway trackbeds and embankments the type of spatial pattern must be determined. To determine this the Poisson and the negative binomial distribution could be used as tools to decide whether the spatial pattern is random or aggregated, respectively.

In addition another way to describe the vegetation is to determine the degree of dispersion using an index of dispersion, the Standardised Morisita Index and the Morisita-Horn index are recommended, since these methods are nearly independent of sample size.

**RQ 3 How reliable are human visual estimates when assessing the extent of vegetation?** After conducting several investigations with both laypersons and domain experts the results show that these are inconsistent in their visual estimates of vegetation (see part IV). The raters were both assessing sample plots on-site on the railway trackbeds/embankments and/or images of the same.

If humans are chosen to assess the extent of vegetation, this investigation highlighted the importance of having a predetermined strict protocol of how to estimate. This would reduce systematic errors made by the misinterpretation of how to assess vegetation cover.

In comparing human visual estimates with the machine vision estimates (see 8), the results show that the human visual estimates were not in agreement with each other. The reliability between human raters were found to be from poor to moderate. It was also found that humans often exaggerated the extent of vegetation cover when compared with the machine vision's output.

**RQ 4 How can the extent of vegetation be measured by making use of machine vision, machine learning, and statistical inference?** In the search for the transfer from manual inspection to machine inspection several experiments were conducted involving machine vision, machine learning and statistical inference. When sampling vegetation along the tracks it was found deemed important to be able to choose the sampling area from the acquired images (or from frames from video clips). Since this allows the sample area to be modified afterwards it enables a more dynamic analysis, i.e. the analysis of images can be redone if desired. To be able to relate to something constant in each image, it was found important to find the tracks in each image. This is because the tracks are the only things in the image which can be assumed to be fairly constant, i.e. the standard track gauge is known to be 1435 mm (plus/minus the tolerance) and the rails are straight lines. After finding the rails, the sampling area could be determined (e.g. along the x-axis, track gauge plus 1 metre, and along y-axis the height of the image/video clip frame). When the sampling area was established the vegetation cover could be computed. See details in chapter 8. The results were evaluated by comparing them with human visual estimates for the sake of validation.

An objective measurement such as the one proposed in this thesis not only offers easy access to the measurements to all the involved parties, but also makes the subcontracting process easier i.e., both the subcontractors and the national railway administration are given the same reference framework concerning vegetation before signing a contract, which can then be cross-checked post maintenance. Biological diversity along the embankments can be mapped, and maintained through better, and robust monitoring procedures. The results achieved in this thesis have shown that the use of image data for detecting vegetation is indeed possible and that such results could form the base for decisions regarding vegetation control.

For the purpose of estimating the vegetation extent, which serves as a base for maintenance decisions, the results are interesting.

The collected raw data in this investigation were the dry weight of plant clusters and individual plants that can be found in areas where the Bilberry and Lichens field layer classes are dominant. Such classes correspond to the Grycksbo and Oxberg data sets, respectively. The collected data were on a high aggregated level. It is important not to step down from this level if we are to draw conclusions about, for example individual plants, or a specific species. These conclusions may be correct, but are only weakly supported by the aggregated data. This erroneous phenomenon is known as the ecological fallacy.

Based on the collected biomass data, the results show that it is possible to use regression to make inferences about the root dry weights by knowing the shoot dry weights. It should be noted that a more extensive study is required to draw more general conclusions, based on the same premises, i.e., to draw conclusions based on collected plant clusters and individual plants, and any type of vegetation.

**RQ 5 How can woody plants growing on railways be recognised using machine vision and machine learning?** One of the goals was to recognise those woody plants that are most likely to grow on railway embankments. Bearing in mind that natural objects are often more heterogeneous within their own class rather than outside it the results present a stable classification performance (see chapter 10). All investigated classifiers recorded a performance of over 84%. The SIFT feature extraction was faster than the other two approaches in particular the MDSIFT. Based on the results the MSDSIFT feature extraction together with the SVM $Chi^2$ kernel classifier exhibited the best performance. One drawback was the long execution time. If time constraints apply, then it might be better to use the DSIFT feature extraction together with the SVM $Chi^2$kernel classifier, which only used 39% of the time used by the MSDSIFT feature extraction together with the SVM $Chi^2$.

**RQ 6 How do measurements using machine vision and machine learning correlate with human visual estimates?** The raters' plotwise mean estimates with the machine vision output, the results show that the human visual estimates do not correlate with the results reported by the machine vision output. That is to say, there was a low degree of association between the human estimates and the MV estimates. As such, the results indicate that it is difficult to fit human estimates by regression with the machine vision result. The results in this investigation were non-significant. Based on the results of these investigations, an automated monitoring approach is suggested, thus transferring the manual inspections into objective monitored inspections by the use of machine vision.

Based on the results of investigations above, an automated monitoring approach is suggested, thus transferring the manual inspections into objective monitored inspections by the use of machine vision.

# Chapter 12

# Future Work

Due to constraints in resources, e.g. time, security, track section availability, and financial constraints, a complete automated prototype for monitoring vegetation on railway trackbeds and embankments was not fully achieved.

Here are some suggestions of where and what to continue working on, by both the scientific community as well as practitioners:

An interesting extension would be to collect data during the hours of darkness. A possible solution would be to highlight the railway embankment with active infrared light, which is not visible to humans. This would overcome such problems as uneven lighting .

Finally it is worth mentioning that although the methodology presented satisfactory results, it should be remembered that the relevant track images acquired in thesis were taken manually. Future work that could solve this kind of problem includes automating the whole setup by placing the camera on a vehicle that is capable of running on the tracks. Such a setup should be aimed at acquiring images or movie-clips automatically, whilst the vehicle is running along the tracks. For such a setup to become reality, however, several key issues need to be addressed, including vehicle and camera positions for image acquisition, vehicle dynamics, vehicle speed and camera response. In the case of movie-clips, a frame sampling procedure could be used.

- The performance of the algorithms developed during the analysis stage should be evaluated. Such a quantitative analysis should incorporate indices, such as means of response times, waiting times, memory loads, queues length, delays, and throughput. This could be solved by simulation, which is not restricted to the few months (May to September) of the year when the vegetation really grows.

- In practice it would be feasible to use a professional camera mounted on the front of a locomotive. The camera should be equipped with a spotmeter where the user can select an off-centre spot, as opposed to the common fixed-centre spot. This calibration procedure should be carried out before each measuring session. The off-centre spot targets and measures a grey card (e.g. QP Card no. 102), which is also mounted on front of a locomotive. The spotmeter will measure and give the exact exposure from time to time based on the standardized colours on the grey card

- It is proposed that each image of the railway embankment be segmented into six areas, as shown in figure 12.1.

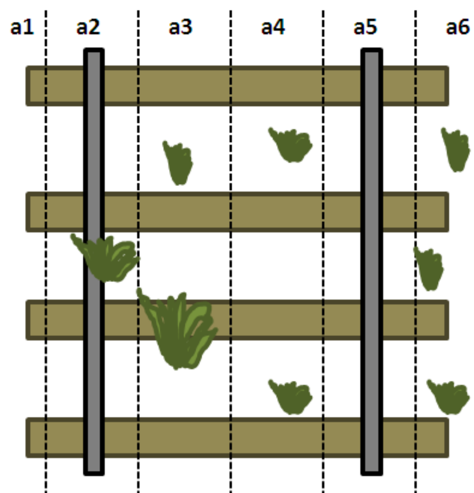



Figure 12.1: Proposed segments of the railway embankment

Seen from above (vertically above the track), these areas really only define spatial boundaries as six intervals along an x-axis, while the y-axis boundaries coincide with the image boundaries. These areas are proposed, firstly, because in terms of rail inspections, some areas (e.g. a2 and a5 in figure 12.1) are considered to be more important than others. This is primarily because, in these two areas, each rail has its connection, or fastening with the sleeper. This proposal was supported by a domain expert (Lundh, J-E., 28 April 2011, Personal interview) at the Swedish Transport Administration. Secondly, it will be possible to detect vegetation that is invading the embankment from the sides, moving in along the x-axis. Thirdly, it would be easier to communicate the proposed segmented areas through images instead of using points and exact coordinates. Each image is already marked with a geographical position that shows the photo point, i.e. the camera's spatial location in longitude, and latitude.

- There is a subset of plants that can quickly form large root systems deep in the railway embankment and are therefore of serious nature. These kinds of species should be identified by the STA. After this has been done, these species could be identified by automated systems (Persson, B., 8 June 2011, Personal interview; Stattin, E., 28 June 2011, Personal interview).

- Classification of the quality of the railway embankment is needed, using a fusion of vegetation feature data and trackbed data, such as macadam, gravel, soil, asphalt, and concrete. The trackbed feature data may include colour shifts, texture, measurements of water content, and measurements of nutrition.

- The influence of background in images is a severe problem, especially when it comes to natural scenes. These often introduce a large number of heterogeneous asymmetric object shapes, textures, and colours. The process of segmenting the background from the foreground (which is the area of interest) can seldom be solved by applying a general algorithm. Algorithms are dependent on the shape, texture, and colours of objects to be identified, or excluded; thus, they keep the remainder, which are the objects of interest. To improve the output of the algorithms, a priori knowledge of the ground condition could be used before scanning a railway section, or area, thus serving as a calibration. The calibration, which implies a choice of algorithm, could reduce the influence of different backgrounds and improve the measurements. For example, if the personnel who are about to scan a railway section know beforehand, or can estimate that most parts of this section will consist of new, grey, sharp ballast it could serve as an a-priori input to the monitoring system.

- Although there is a correlation between biomass and the cover attribute, it would add more strength to the analysis if a third dimension namely the height of the vegetation. This could be accomplished by optical laser distance measurement sensors. These sensors typically measures the distance to objects by pulsing out laser beams on an area. The result is a three dimensional point cloud compromising a surface. The height of the vegetation, together with the cover area (the mentioned 3D surface), would give information about the individual plant volume, or plant cluster volume. Based on an estimate of the density (i.e. weight unit per volume unit) of a plant, the biomass could then be estimated. Investigations in area have already been initiated.

- Investigate in the winter months as whether it is feasible to find woody plants, especially when a thin layer of snow has fallen, covering the ballast and most of the low layers of graminoids and herbs. This would facilitate the image segmentation process. The optimal layer thickness has yet of course to be discovered, but would probably have to be in the interval of 1 to 10 cm thick.

- Investigate the spatial distribution of plants growing on railway embankments by testing the hypothesis that they follow a clumped (clustered) distribution Mauseth (1998, p.733).

- An extended field experiment like the one carried out on the Älvdalsbanan in Oxberg should be conducted as soon as possible. The purpose is to gather a large amount of data on woody plants that grow on railways in this as well as other climate zones.

- Acquire more rater data relating to the estimation of cover from trackbeds onsite and images (see chapter 5) and try to engage more qualified raters, i.e. domain experts.

- Further investigate whether it is feasible to monitor the ballast and/or its surroundings in order to characterise the soil. Personal conversations with domain experts during a workshop [Workshop at SLU, Alnarp 27-28 April 2011] indicate that this is viable . One application could be to classify the field layer types shown in table 5.1.

- Investigate remote sensing as a means of monitoring the Rights-of-Way corridor, which runs 20 metres either side of the track. A pilot study has already been launched in the form of a master of science thesis study, with the author of this thesis as the supervisor.

- Investigate the quality of the STA STRIX wagon video-clips captured by cameras that view the front of the wagon.

- To test the results of vegetation cover and the root to shoot analyses, a quantitative sampling study should be conducted; for example, it could take the form of a series of systematic samplings along railways.

- In addition to monitoring vegetation by use of machine vision, also investigate the feasibility of locating objects like rail fastenings, and base-plates. The purpose of this is to enable other ways of identifying the location of the edge of sleepers, particularly if some sleepers cannot be found. This also enables the monitoring system to be used in activities other than vegetation monitoring.

# References


Alchanatis V, Ridel L, Hetzroni A, Yaroslavsky L (2005) Weed detection in multi-spectral images of cotton fields. *Computers and Electronics in Agriculture* 47:243 – 260.

Alexander M (2008) Survey, surveillance, monitoring and recording In *Management Planning for Nature Conservation*, pp. 49–62. Springer Netherlands, 2008 edition.

Altman DG, Bland JM (1983) Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician)* 32:pp. 307–317.

Anscombe FJ (1950) Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika* pp. 358–382.

Banverket (1995) Bvf 586.10 - skarvfritt spar. regler for byggande och underhall National Railway Regulation.

Banverket (1996) Bvh 586.40 - spårteknik. spårgeometrihandboken National Railway Manual.

Banverket (2000) Bvh 827.1 - handbok om vegetation National Railway Manual.

Banverket (2001) Bvh 827.2 - behovsanalys infor vegetationsreglering National Railway Manual.

Banverket (2005a) Bvh 807.30 - underhallsbesiktning av banoverbyggnad National Railway Manual.

Banverket (2005b) Bvh 807.31 - underhallsbesiktning av banunderbyggnad National Railway Manual.

Banverket (2005c) Bvh 807.33 - underhallsbesiktning av elkraftanlaggningar National Railway Manual - Maintenance inspection of electrical power plants.

Banverket (2005d) Bvh 807.34 underhallsbesiktning av signalanlaggningar National Railway Manual.

Banverket (2005e) Bvh 807.36 - underhallsbesiktning av ovriga anlaggningar National Railway Manual.

Barker PCJ (2001) *A technical manual for vegetation monitoring / Phil Barker* Resource Management and Conservation, Dept. of Primary Industries, Water and Environment, Hobart, Tas. : BibTex package created from National Library of Australia Catalogue http://catalogue.nla.gov.au.

Barnett V (2003) *Sample Survey Principles and Methods* A Hodder Arnold Publication. John Wiley & Sons, 3 rev ed edition.

Below MDBA, Gachter FSA, Kuppelwieser HSA (2003) Uic vegetation control project final report Technical report ISBN 2-7461-0744-9, International Union of Railways (UIC), Paris.

Benavides JC, Jesús ISD (2009) Digitized images provide more accuracy and efficiency to estimate bryophyte cover. *The Bryologist* 112:pp. 12–18.

Bishop CM (2006) *Pattern recognition and machine learning*, Vol. 4 Springer New York.

Bland JM, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet* 327:307 – 310 Originally published as Volume 1, Issue 8476.

Bonham C (1989) *Measurements for terrestrial vegetation* Wiley Interscience publication. Wiley.

Bonham CD, Clark DL (2005) Quantification of plant cover estimates. *Grassland Science* 51:129–137.

Bosch A, Zisserman A, Muoz X (2007) Image classification using random forests and ferns In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–8. IEEE.

Brakenhielm S, Qinghong L (1995) Comparison of field methods in vegetation monitoring. *Water, Air, & Soil Pollution* 79:75–87 10.1007/BF01100431.

Braun-Blanquet J (1932) *Plant Sociology - The Study of Plant Communities* McGraw-Hill Book Company, Inc. New York and London Authorized English Translation of Pflanzensoziologie.

Briggs K (2010) Charing embankment: climate change impacts on embankment hydrology. *Ground Engineering* 43:28–31.

BSI (2010) Bs en 13306:2010 maintenance - maintenance terminology.

Campbell J, Wynne R (2011) *Introduction to Remote Sensing, Fifth Edition* Guilford Publications.

Caragounis V (2014) Handbok för ras-, skred- och slamströmsolycka Manual, Myndigheten för samhällsskydd och beredskap (MSB) Eng. Manual to landslides, mudslides and debris flow disaster by the Swedish Civil Contingencies Agency.

Cardina J, Johnson GA, Sparrow DH (1997) The nature and consequence of weed spatial distribution. *Weed Science* 45:364–373.

Cardina J, Sparrow DH, McCoy EL (1995) Analysis of spatial distribution of common lambsquarters (chenopodium album) in no-till soybean (glycine max). *Weed Science* 43:pp. 258–268.

Cavalier-Smith T (2004) Only six kingdoms of life. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271:1251–1262.

Cederlund H, Fogelberg F, Hansson D, Nyberg RG, Schroeder H (2014) Utveckling av metod for att bedoma behovet av ograsbekampning i spar Technical report, Dalarna University, Computer Engineering Externa finansiarer: Trafikverket.

Chandra S et al. (2008) *Railway engineering* Oxford University Press, Inc.

Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment* 6:284–290.

Cochran WG (2007) *Sampling techniques* John Wiley & Sons.

Cohen J (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20:37–46.

Cohen J (1968) Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* .

Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24:603–619.

Coulloudon B, Eshelman K, Gianola J, Habich Nea (1999) Sampling vegetation attributes Interagency Technical Reference BLM/RS/ST-96/002+1730, Bureau of Land Management's National Applied Resource Sciences Center, BC-650B. P.O. Box 25047. Denver, Colorado 80225-0047.

Cox D, Reid N (2000) *The Theory of the Design of Experiments* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC.

Csurka G, Dance C, Fan L, Willamowski J, Bray C (2004).

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, Vol. 1, pp. 886–893 vol. 1.

Dallal GE (2012) The Little Handbook of Statistical Practice Internet as Kindle version. 1.10.

Daubenmire R (1959) A canopy coverage method of vegetation analysis. *Northwest Science* 33:43–64.

DeCarlo LT (1997) On the meaning and use of kurtosis. *Psychological methods, American Psychological Association* 2(3):pp. 292–307.

Dept. of Soil and Environment at SLU (2012) Vegetation types Internet Reprint permission via mail by prof. Lars Lundin, SLU (2014-12-03).

Dougherty M (2013) Sound and vision. *Journal of Intelligent Systems* 22:213–214.

Duda RO, Hart PE, Stork DG (2000) *Pattern Classification* Wiley-Interscience, 2 edition.

Eddowes M, Waller D, Taylor P, Briggs B, Meade T, Ferguson I (2003) Railway safety implications of weather, climate and climate change: Final report to the railway safety and standards board Technical report Issue 2, AEA (Atomic Energy Authority) Technology Plc, Didcot, UK. Report nr: AEAT/RAIR/76148/R03/005.

Elzinga C, Salzer D, Willoughby J, J.P.Gibbs (2001) *Monitoring Plant and Animal Populations* Blackwell Science.

Elzinga C, Salzer D, Willoughbyh J (1998) Measuring and monitoring plant populations Technical Reference 1730-1 BLM/RS/ST-98/005+1730 (5.1 MB), Bureau of Land Management. Denver, Colorado. USDI, BLM.

Fehmi JS (2010) Confusion among three common plant cover definitions may result in data unsuited for comparison. *Journal of Vegetation Science* 21:273–279.

Field A, Miles J, Field Z (2012) *Discovering Statistics Using R* SAGE Publications.

Fischler M, Firschein O (1987) *Intelligence: The Eye, the Brain, and the Computer* Addison-Wesley.

Fleiss JL (1971) Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76:378–382.

Floyd DA, Anderson JE (1987) A comparison of three methods for estimating plant cover. *Journal of Ecology* 75:pp. 221–228.

Flusser J, Zitova B, Suk T (2009) *Moments and Moment Invariants in Pattern Recognition* Wiley Publishing.

Fogelberg F (2011) Teknik for bedomning av vegetationsbekampningsbehovet pa banvallar . en forstudie Uppdragsrapport, JTI (Institutet for jordbruks och miljoteknik), Uppsala, Sweden Presented at 9th EWRS Workshop on Physical and Cultural Weed Control Samsun, Turkey, 28 - 30 March 2011. Maintenance of railway embankments how to use the Yara N-sensorfor prognosis of weed control measures.

Fontana A, Frey JH (1994) *Interviewing: The Art of Science in The Handbook of Qualitative Research* Thousand Oaks: Sage Publications.

Freedman D (2001) *Ecological inference and the ecological fallacy in International Encyclopedia of the Social & Behavioral Sciences*, Vol. 6 Elsevier.

Fukunaga K, Hostetler L (1975) The estimation of the gradient of a density function, with applications in pattern recognition. *Information Theory, IEEE Transactions on* 21:32–40.

Gardener M (2012) *Statistics for Ecologists Using R and Excel: Data Collection, Exploration, Analysis and Presentation* Pelagic Publishing.

Ghahramani Z (2004) *Advanced Lectures in Machine Learning. Lecture Notes in Computer Science*, Vol. 3176 Springer Science & Business Media Unsupervised learning p. 72–112.

Goldsmith F (2012) *Monitoring for Conservation and Ecology* Conservation Biology. Springer Netherlands.

Gonzalez RC, Woods RE (2007) *Digital Image Processing (3rd Edition)* Prentice Hall.

Goodall D (1952) Some considerations in the use of point quadrats for the analysis of vegetation. *Australian Journal of Biological Sciences* 5:1 – 41.

Greig-Smith P (1983) *Quantitative Plant Ecology*, Vol. 9 of Studies in ecology University of California Press, Berkeley, 3 ed. edition.

Hagglund B, Lundmark JE (1987) *Handledning i bonitering med Skogshogskolans boniteringssystem. D. 1, Definitioner och anvisningar* Skogsstyr., Jonkoping, 3. tr. edition.

Hagglund B, Lundmark JE (2003) *Handledning i bonitering med Skogshogskolans boniteringssystem. D. 2, Diagram och tabeller* Skogsstyr., Jonkoping, 4. uppl. edition.

Hagglund B, Lundmark JE (2004) *Handledning i bonitering med Skogshogskolans boniteringssystem. D. 3, Markvegetationstyper, skogsmarksflora* Skogsstyrelsen, Jonkoping, 4. uppl. ; 5. uppl., 2010 edition.

Hallgren KA (2012) Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in quantitative methods for psychology* 8:23.

Hayakawa T (2007) Vegetation control through selective elimination of large-size weeds on embankments using growth retardant. *Quarterly Report of RTRI (Railway Technical Research Institute)* Vol. 48:240–243.

Hayes AF, Krippendorff K (2007) Answering the call for a standard reliability measure for coding data. *Communication methods and measures* 1:77–89.

Haykin S (2009) *Neural Networks and Learning Machines* Pearson International Edition. Pearson.

Heisel T, Christensen S (1998) A digital camera system for weed detection. *Proceedings of the 4th international conference on precision agriculture* July:p1569–1577.

Hill D (2005) *Handbook of biodiversity methods: survey, evaluation and monitoring* Cambridge University Press.

Horn HS (1966) Measurement of "overlap" in comparative ecological studies. *American naturalist* Vol. 100:419–424.

Huisman M (2001) Reglering av vedartad vegetation utmed jarnvagar och vagar - en litteraturstudie over kunskapslaget (control of brush and trees in rights-of-way - a literature study) Technical report, Swedish University of Agricultural Sciences Department of Agricultural Engineering.

Hunt R, Lloyd PS (1987) Growth and partitioning. *New Phytologist* 106:pp. 235–249.

Hurford C (2006) *Monitoring Nature Conservation in Cultural Habitats: A Practical Guide And Case Studies* SpringerLink: Springer e-Books. Springer London, Limited.

Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:pp. 187–211.

Huxley JS, Teissier G (1936) Terminology of relative growth. *Nature* 137:780–781.

J. Yang KY, Gong Y, Huang T (2009) Linear spatial pyramid matching using sparse coding for image classification In *Proc. of CVPR'09*.

Jardine AK, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech-*

*anical Systems and Signal Processing* 20:1483–1510.

Jensen J (2007) *Remote sensing of the environment: an earth resource perspective* Prentice Hall series in geographic information science. Pearson Prentice Hall.

Joachims T (1998) *Text categorization with support vector machines: Learning with many relevant features* Springer.

Jonasson S (1988) Evaluation of the point intercept method for the estimation of plant biomass. *Oikos* 52:pp. 101–106.

Jukola-Sulonen EL, Salemaa M (1985) A comparison of different sampling methods of quantitative vegetation analysis. *Silva Fennica* 19:325–337.

Khazraei K, Deuse J (2011) A strategic standpoint on maintenance taxonomy. *Journal of Facilities Management* Vol. 9 Iss: 2:96–113.

Klimes L (2003) Scale-dependent variation in visual estimates of grassland plant cover. *Journal of Vegetation Science* 14:815–821.

Krebs CJ, Danell K, Angerbjorn A, Agrell J, Berteaux D, Brathen KA, Danell O, Erlinge S, Fedorov V, Fredga K, Hjalten J, Hogstedt G, Jansdottir IS, Kenney AJ, Kjellen N, Nordin T, Roininen H, Svensson M, Tannerfeldt M, Wiklund C (2003) Terrestrial trophic dynamics in the canadian arctic. *Canadian Journal of Zoology* 81:827–843.

Krebs C (1999) *Ecological Methodology* The Benjamin/Cummings series in the life sciences. Benjamin/Cummings.

Krippendorff K (1992) Recent developments in reliability analysis In *Annual Meeting of the International Communication Assoc.*, Miami, FL.

Krippendorff K (2004) *Content Analysis: An Introduction to Its Methodology* Sage.

Kuchler AW (1966) Analyzing the physiognomy and structure of vegetation. *Annals of the Association of American Geographers* 56:pp. 112–127.

Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, Vol. 2, pp. 2169–2178.

Ledig FT, Bormann FH, Wenger KF (1970) The distribution of dry matter growth between shoot and roots in loblolly pine. *Botanical Gazette* 131:349–359.

Lee J, Abujamra R, Jardine AK, Lin D, Banjevic D (2004) An integrated platform for diagnostics, prognostics and maintenance optimization

In *Proceedings of the Intelligent Maintenance Systems 2004*.

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60:91–110.

Manel S, Williams HC, Ormerod SJ (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of applied Ecology* 38:921–931.

Mannetje L, Jones R (2000) *Field and Laboratory Methods for Grassland and Animal Production Research* Cabi Series. CABI Pub.

March ST, Smith GF (1995) Design and natural science research on information technology. *Decis. Support Syst.* 15:251–266.

Martin P, Bateson P (2007) *Measuring Behaviour: An Introductory Guide* Cambridge University Press.

Mauseth J (1998) *Botany an introduction to plant biology* Jones and Bartlett.

McDonald J (2014) *Handbook of Biological Statistics* Sparky House Publishing, Baltimore, Maryland, 3 edition.

McGill R, Tukey JW, Larsen WA (1978) Variations of box plots. *The American Statistician* 32:pp. 12–16.

McIntosh Distribution (2015) Weedseeker Internet http://www.mcintoshdistribution.com.au (Retrieved 2 Apr. 2015).

Monk C (1966) Ecological importance of root/shoot ratios. *Bulletin of the Torrey Botanical Club* 93:pp. 402–406.

Morisita M (1959) Measuring of interspecific association and similarity between communities. *Mem. Fac. Sci. Kyushu Univ. Series E* 3:65–80.

Mueller-Dombois D, Ellenberg H (1974) *Aims and methods of vegetation ecology* Blackburn Press, 2003 edition.

Nationalencyklopedin (2011) Tall Retrieved 08 Aug 2011.

Natrella M, Croarkin C, Tobias P (2012) NIST/SEMATECH e-Handbook of Statistical Methods Internet Acc. 2014-08-10.

Niklas KJ (2005) Modelling below- and above-ground biomass for non-woody and woody plants. *Annals of Botany* 95:315–321.

Nixon M, Aguado AS (2008) *Feature Extraction & Image Processing, Second Edition* Academic Press, 2 edition.

Nowak E, Jurie F, Triggs B (2006) Sampling strategies for bag-of-features image classification In *Computer Vision–ECCV 2006*, pp. 490–503. Springer.

Nyberg R, Gupta N, Yella S, Dougherty M (2014) Machine vision for condition monitoring vegetation on railway embankments In *Railway Condition Monitoring (RCM 2014), 6th IET Conference on*, pp. 1–7.

Nyberg RG, Gupta N, Yella S, Dougherty M (2013a) Detecting plants on railway embankments. *Journal of Software Engineering and Applications* 6:8–12.

Nyberg RG, Gupta NK, Yella S, Dougherty MS (2013b) Monitoring vegetation on railway embankments: Supporting maintenance decisions In *Proceedings of the Int. Conference on Ecology and Transportation*.

Oates BJ (2006) *Researching Information Systems and Computing* Sage Publications Ltd.

O'Hara S, Draper BA (2011) Introduction to the bag of features paradigm for image classification and retrieval. *arXiv preprint arXiv:1101.3354* .

Otsu N (1979) A threshold selection method from gray-level histograms. *Systems, Man and Cybernetics, IEEE Transactions on* 9:62 –66.

Pearsall WH (1927) Growth studies: Vi. on the relative sizes of growing plant organs. *Annals of Botany* os-41:549–556.

Perez AJ, Lopez F, Benlloch JV, Christensen S (1997) Colour and shape analysis techniques for weed detection in cereal fields. *First European Conference for Information Technology in Agriculture, Copenhagen* .

Philipp I, Rath T (2002) Improving plant discrimination in image processing by use of different colour space transformations. *Computers and Electronics in Agriculture* 35:1 – 15.

Prince S (2012) *Computer Vision: Models Learning and Inference* Cambridge University Press.

Profillidis V (2006) *Railway management and engineering* Ashgate.

R Core Team (2014) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0.

Railway Accident Investigation Unit R (2010) Road vehicle struck at level crossing xm096, county roscommon, 2nd september 2010 Accident investigation report Investigation Report 2011-R006, RAIU.

Rankin G, Stokes M (1998) Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation* 12:187–199.

Rottgermann M, Steinlein T, Beyschlag W, Dietz H (2000) Linear relationships between aboveground biomass and plant cover in low open herbaceous vegetation. *Journal of Vegetation Science* 11:145–148.

Russell B, Benner C, Wischkaemper J, Jewell W, McCalley J (2007) Reliability based vegetation management through intelligent system monitoring Technical report, Power Systems Engineering Research Center.

Sargent C (1984) *Britain's railway vegetation* Institute of Terrestrial Ecology.

Scott J, Loveridge F, O'Brien A (2007) Influence of climate and vegetation on railway embankments In *XIV European Conference on Soil Mechanics and Geotechnical Engineering, Madrid, ES*.

Scott WA (1955) Reliability of content analysis:: The case of nominal scale coding. *Public Opinion Quarterly* 19:321–325.

Scrivner R, editor (2004) *Railroad Vegetation Management*, Vol. 56. Monsanto, Proceeding of the California Weed Science Society.

SFNI SNFI (2011) Standing volume for different tree species by diameter class, 2007-2011.

Shingleton A (2010) Allometry: The study of biological scaling. nature education knowledge 1(9):2 www.

Shoukri MM (2003) *Measures of interobserver agreement and reliability* CRC Press.

Shrout P, Fleiss J (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin* 86:420–428.

Siegel S (1957) Nonparametric statistics. *The American Statistician* 11:pp. 13–19.

Silva MB, Hermo BS, García-Rodeja E, Freire NV (2005) Reutilization of granite powder as an amendment and fertilizer for acid soils. *Chemosphere* 61:993 – 1002.

Simon HA (1996) *The sciences of the artificial* MIT Press, Cambridge, MA, USA, 3 edition.

Sivic J, Zisserman A (2003) Video google: A text retrieval approach to object matching in videos pp. 1470–1477 vol.2.

Smethurst J, Clarke D, Powrie W (2012) Factors controlling the seasonal variation in soil water content and pore water pressures within a lightly vegetated clay slope. *Geotechnique* 62:429–446.

Sokal RR, Rohlf J (2011) *Biometry* W. H. Freeman.

Stevens S (1946) On the theory of scales of measurement. *Science* 103:677–680.

Swedish Forest Soil Inventory (2014) Markinventeringen Internet.

Swedish University of Agricultural Sciences (2011) Swedish national forest

inventory Internet (Retrieved 2 July 2011).
Torstensson L (2001) Use of herbicides on railway tracks in sweden. *Pesticide Outlook* 12:pp. 16–21.
Trafikverket (2012a) Bvf 807.2 - sakerhetsbesiktning av fasta jarnvagsanlaggningar National Railway Regulation.
Trafikverket (2012b) Bvs 1585.005 typsektioner för banan National Railway Standard. Reference No. TRV 2012/38092.
Trafikverket (2012c) Bvs 1586.20 banoverbyggnad - infrastrukturprofiler National Railway Regulation, Reference No. TRV 2012/71445.
Trafikverket (Swedish Transport Administration) (2015) How we procure Internet Acc. 2014-08-30.
Tsai CF (2012) Bag-of-words representation in image annotation a review. *ISRN Artificial Intelligence* 2012:19.
Tukey J (1977) *Exploratory Data Analysis* Addison-Wesley series in behavioral science. Addison-Wesley Publishing Company.
Unisys Weather (2012) Synop data format (fm-12) surface synoptic observations Internet Retrieved 8 Aug. 2012.
Vrindts E, De Baerdemaeker J (1996) Feasibility of weed detection with optical reflection measurement. *Proceedings of the Brighton Crop Protection Conference - Pest and Diseases* 1:443–444.
Wang J, Yang J, Yu K, Lv F, Huang T, Gong Y (2010) Locality-constrained linear coding for image classification In *CVPR*.
Wang M, Vandermaar A, Srivastava K (2002) Review of condition assessment of power transformers in service. *Electrical Insulation Magazine, IEEE* 18:12–25.
Wang N, Zhang N, Wei J, Stoll Q, Peterson D (2007) A real-time, embedded, weed-detection system for use in wheat fields. *Biosystems Engineering* 98:276 – 285.
Wartenberg G, Schmidt H (1999) Optoelectronic sensors - an alternative for weed detection. *Landtechnik* 54 (6):340–341.
Weir JP (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. *The Journal of Strength & Conditioning Research* 19:231–240.
Weiss R (1994) *Learning from Strangers: The Art and Method of Qualitative Interview Studies* Free Press.
Whittaker RH (1975) *Communities and Ecosystems* MacMillan, New York.
Wilson JB, Bastow J (1988) A review of evidence on the control of shoot: Root ratio, in relation to models. *Annals of Botany* 61:433–449.

Wolda H (1981) Similarity indices, sample size and diversity. *Oecologia* 50:296–302.

Yam RCM, Tse P, Li L, Tu P (2001) Intelligent predictive decision support system for condition-based maintenance. *The International Journal of Advanced Manufacturing Technology* 17:383 – 391.

Yin RK (2014) *Case study research: Design and methods* Sage publications, 5th ed. edition 312 pages.

# Interviews


Below, Dr. Michael [Deutsche Bahn AG, Germany] at the Workshop *Sustainable management of rail environment*, Alnarp Sweden. 27-28 April 2011*. Personal interview.*

Lindström, Anders. [Professor of Science in Forestry, and head of the department of Forest and Wood Technology at Dalarna University, Sweden.] 07 September 2011. *Personal interview.*

Lundh, Jan-Erik. [Licentiate of Science in Forestry, and vegetation Engineer at the Swedish Transport Administration, Sweden.]. 28 April 2011. *Personal interview.*

Persson, Bengt [Docent of Forest Husbandry at Dalarna University, Sweden.] 08 June 2011. *Personal interview.*

Sjöblom, Tomas [ CEO at Scandinavian Track Group AB, Sweden ] 01 March 2011. *Personal interview*.

Stattin, Eva [Docent of Science in Forestry at Dalarna University, Sweden] 28 June 2011. *Personal interview.*

# Appendix A

# Data from Falun - Grycksbo Railway, Sweden

## A.1 Plant Root and Shoot Dry Weights

The root and shoot dry weights in grams are presented in table A.1. Chapter 5 describes how the data were acquired. Section 9 describes the analysis of the data.

| plotID | No. of individ./ clusters | Specie | Root dry weight | Shoot dry weight |
|---|---|---|---|---|
| 1 | 1 | Solidago virgaurea | 11.9767 | 2.2717 |
| 1 | 1 | Ranunculus acris | 7.0337 | 38.8937 |
| 1 | 21 | Taraxacum Sp. | 22.3237 | 16.2127 |
| 1 | 1 | Hieracium sp. | 5.4757 | 5.8307 |
| 2 | 1 | Betula pubescens | 2.16015 | 1.61915 |
| 2 | 1 | Veronica serpyllifolia | 1.51015 | 0.23215 |
| 2 | 1 | Tanacetum vulgare | 0.29215 | 0.05915 |
| 2 | 1 | Heracleum sphondylium L. | 6.59415 | 5.75815 |
| 2 | 9 | Solidago virgaurea | 31.1547 | 19.3117 |
| 2 | 23 | Taraxacum Sp. | 19.1057 | 11.8017 |
| 2 | 4 | Deschampsia Flexuosa | 53.5737 | 10.1167 |
| 3 | 20 | Solidago virgaurea | 69.6537 | 37.6047 |
| 3 | 1 | Betula pubescens | 3.16315 | 3.06815 |
| 3 | 15 | Taraxacum Sp. | 3.13115 | 4.53115 |
| 3 | 2 | Heracleum sphondylium L. | 11.01015 | 6.20215 |
| 3 | 1 | Alnus incana | 4.07715 | 8.77315 |
| 3 | 1 | Pinus sylvestris | 0.45115 | 0.13715 |
| 3 | 7 | Hieracium sp. | 1.36015 | 1.21015 |
| 3 | 3 | Deschampsia Flexuosa | 1.75615 | 0.85915 |
| 4 | 23 | Hieracium sp. | 12.8421 | 17.0391 |
| 4 | 6 | Taraxacum Sp. | 7.61815 | 7.84415 |
| 4 | 1 | Festuca L. | 1.11415 | 0.53815 |
| 5 | 12 | Hieracium sp. | 8.9001 | 16.6701 |
| 5 | 17 | Taraxacum Sp. | 9.7181 | 8.6321 |
| 5 | 1 | Heracleum sphondylium L. | 5.81915 | 2.36315 |
| 5 | 1 | Deschampsia Flexuosa | 0.56415 | 0.45615 |

Table A.1: Root and shoot dry weights (in grams) collected at plot 1 to 5 (Grycksbo, Sweden)

# Appendix B

# Data from Älvdalsbanan, Oxberg, Sweden

## B.1 Plant Root and Shoot Dry Weights

The root dry weights and shoot dry weights from around 300 plants taken from the sample plots 200 to 204 on the railway between Mora and Märbäck are presented in tables B.1, B.2, B.3 and B.4. The measured unit is grams. Chapter 5 describes how the data were acquired and section 9 describes the analysis of the data. The number of plant individuals and/or plant clusters is approximate. This was because of the difficulty of defining individual plants. This was especially the case with graminoids (grass), where the ”number of” column was often left blank.

Figure B.1 shows the sample grid frame placed on each of the five sample plots, 200 to 204.


Figure B.1: The five sample plot areas

| plotID | Specie | No. of individ./ clusters | Root dry weight | Shoot dry weight |
|---|---|---|---|---|
| 200 | Betula pendula | 2 | 0.2214 | 0.3144 |
| 200 | Festuca L. | | 55.04666667 | 55.84666667 |
| 200 | Festuca L. | | 85.17666667 | 87.19666667 |
| 200 | Picea abies | 4 | 1.753666667 | 3.014666667 |
| 200 | Picea abies | 1 | 1.568666667 | 8.473666667 |
| 200 | Picea abies | 3 | 2.118666667 | 5.552666667 |
| 200 | Picea abies | 5 | 0.4774 | 1.2054 |
| 200 | Picea abies | 2 | 0.6694 | 1.6514 |
| 200 | Picea abies | 1 | 41.6555 | 127.6555 |
| 200 | Picea abies | 1 | 39.6555 | 100.1555 |
| 200 | Picea abies | 1 | 15.4555 | 57.4555 |
| 200 | Pinus sylvestris | 2 | 1.214666667 | 5.529666667 |
| 200 | Pinus sylvestris | 2 | 1.122666667 | 5.698666667 |
| 200 | Pinus sylvestris | 1 | 1.811666667 | 8.907666667 |
| 200 | Pinus sylvestris | 2 | 15.919 | 18.415 |
| 200 | Pinus sylvestris | 2 | 16.372 | 23.318 |
| 200 | Pinus sylvestris | 1 | 0.1854 | 0.6844 |
| 200 | Pinus sylvestris | 1 | 0.0424 | 0.3834 |
| 200 | Rubus idaeus | 1 | 1.2914 | 0.9274 |
| 201 | Betula pubescens | 2 | 1.126666667 | 1.581666667 |
| 201 | Betula pubescens | 4 | 0.1584 | 0.1694 |
| 201 | Betula pubescens | 2 | 0.2354 | 0.1854 |
| 201 | Betula pubescens | 1 | 0.2514 | 0.1574 |
| 201 | Festuca L. | | 29.34666667 | 48.92666667 |
| 201 | Picea abies | 5 | 3.266666667 | 8.456666667 |
| 201 | Picea abies | 2 | 2.301666667 | 4.775666667 |

Table B.1: Root and shoot dry weights (in grams) collected from plots 200 to 201 (Oxberg, Sweden)

| plotID | Specie | No. of individ./ clusters | Root dry weight | Shoot dry weight |
|---|---|---|---|---|
| 201 | Picea abies | 11 | 3.631666667 | 9.739666667 |
| 201 | Picea abies | 6 | 1.753666667 | 2.786666667 |
| 201 | Picea abies | 2 | 0.7594 | 1.4854 |
| 201 | Pinus sylvestris | 4 | 3.698666667 | 23.45066667 |
| 201 | Pinus sylvestris | 6 | 2.906666667 | 18.41966667 |
| 201 | Pinus sylvestris | 13 | 2.071666667 | 8.742666667 |
| 201 | Pinus sylvestris | 4 | 0.3004 | 0.7664 |
| 201 | Pinus sylvestris | 2 | 0.6954 | 4.0794 |
| 201 | Pinus sylvestris | 5 | 2.011666667 | 11.03066667 |
| 202 | Festuca L. | 1 | 0.4584 | 1.2644 |
| 202 | Festuca L. | 1 | 1.3494 | 3.7214 |
| 202 | Festuca L. | 1 | 1.2004 | 2.7544 |
| 202 | Festuca L. | 1 | 0.7444 | 1.3024 |
| 202 | Picea abies | 1 | 7.362666667 | 21.87966667 |
| 202 | Picea abies | 9 | 1.455666667 | 2.225666667 |
| 202 | Picea abies | 1 | 1.648666667 | 3.617666667 |
| 202 | Picea abies | 5 | 1.409666667 | 2.798666667 |
| 202 | Pinus sylvestris | 10 | 2.129666667 | 10.58166667 |
| 202 | Pinus sylvestris | 1 | 1.784666667 | 10.12066667 |
| 202 | Pinus sylvestris | 6 | 1.968666667 | 12.19866667 |
| 202 | Pinus sylvestris | 9 | 1.486666667 | 2.830666667 |
| 202 | Pinus sylvestris | 2 | 0.1044 | 0.1704 |
| 203 | Betula pendula | 1 | 0.1964 | 0.533666667 |
| 203 | Betula pubescens | 1 | 0.1914 | 0.2504 |
| 203 | Betula pubescens | 1 | 0.0294 | 0.0884 |
| 203 | Betula pubescens | 3 | 0.1384 | 0.2824 |

Table B.2: Root and shoot dry weights (in grams) collected from plots 201 to 203 (Oxberg, Sweden)

| plotID | Specie | No. of individ./ clusters | Root dry weight | Shoot dry weight |
|---|---|---|---|---|
| 203 | Betula pubescens | 4 | 0.0794 | 0.1394 |
| 203 | Festuca L. | | 2.3254 | 0.1754 |
| 203 | Festuca L. | | 2.8264 | 3.0704 |
| 203 | Festuca L. | | 1.4814 | 1.8784 |
| 203 | Festuca L. | | 1.3264 | 1.4364 |
| 203 | Festuca L. | | 0.5344 | 1.4554 |
| 203 | Festuca L. | | 0.7604 | 1.8374 |
| 203 | Festuca L. | | 0.4444 | 0.3454 |
| 203 | Festuca L. | | 0.4734 | 0.8114 |
| 203 | Festuca L. | | 0.2324 | 0.4924 |
| 203 | Picea abies | 1 | 5.118666667 | 8.007666667 |
| 203 | Picea abies | 3 | 0.5264 | 1.576666667 |
| 203 | Picea abies | 5 | 0.6764 | 1.532666667 |
| 203 | Picea abies | 4 | 0.1604 | 0.492666667 |
| 203 | Picea abies | 1 | 0.1754 | 1.576666667 |
| 203 | Pinus sylvestris | 8 | 1.740666667 | 7.114666667 |
| 203 | Pinus sylvestris | 10 | 1.046666667 | 3.021666667 |
| 203 | Pinus sylvestris | 10 | 2.354666667 | 11.27966667 |
| 203 | Pinus sylvestris | 1 | 0.244666667 | 1.645666667 |
| 203 | Pinus sylvestris | 6 | 0.747666667 | 2.804666667 |
| 203 | Pinus sylvestris | 7 | 1.016666667 | 3.992666667 |
| 203 | Pinus sylvestris | 1 | 0.705666667 | 5.199666667 |
| 203 | Pinus sylvestris | 2 | 1.272666667 | 5.360666667 |
| 203 | Pinus sylvestris | 1 | 0.4234 | 1.812666667 |
| 204 | Festuca L. | 1 | 0.3654 | 0.6974 |
| 204 | Festuca L. | 1 | 1.6964 | 0.9754 |

Table B.3: Root and shoot dry weights (in grams) collected from plots 203 to 204 (Oxberg, Sweden)

| plotID | Specie | No. of individ./ clusters | Root dry weight | Shoot dry weight |
|---|---|---|---|---|
| 204 | Festuca L. | 1 | 4.3014 | 2.6314 |
| 204 | Festuca L. | 1 | 0.6774 | 1.0994 |
| 204 | Festuca L. | 1 | 1.0774 | 1.8604 |
| 204 | Festuca L. | 1 | 1.3364 | 1.5564 |
| 204 | Festuca L. | 1 | 0.2594 | 0.7364 |
| 204 | Festuca L. | 1 | 2.3004 | 3.0544 |
| 204 | Festuca L. | 1 | 0.6314 | 0.9554 |
| 204 | Picea abies | 1 | 0.0714 | 0.3324 |
| 204 | Picea abies | 1 | 0.1944 | 0.3454 |
| 204 | Picea abies | 1 | 0.1284 | 0.3864 |
| 204 | Picea abies | 1 | 0.1794 | 0.2334 |
| 204 | Picea abies | 1 | 0.2744 | 0.6304 |
| 204 | Picea abies | 1 | 0.0444 | 0.4084 |
| 204 | Pinus sylvestris | 1 | 0.5244 | 4.042666667 |
| 204 | Pinus sylvestris | 2 | 0.2974 | 1.645666667 |
| 204 | Pinus sylvestris | 1 | 0.1134 | 1.481666667 |
| 204 | Pinus sylvestris | 1 | 0.2814 | 2.071666667 |
| 204 | Pinus sylvestris | 1 | 0.1934 | 1.616666667 |
| 204 | Pinus sylvestris | 1 | 0.1604 | 1.050666667 |
| 204 | Pinus sylvestris | 3 | 0.5074 | 2.657666667 |
| 204 | Pinus sylvestris | 2 | 0.1074 | 1.427666667 |
| 204 | Pinus sylvestris | 4 | 0.1664 | 1.040666667 |
| 204 | Pinus sylvestris | 1 | 0.2854 | 1.279666667 |
| 204 | Pinus sylvestris | 1 | 0.0954 | 1.343666667 |
| 204 | Pinus sylvestris | 1 | 0.0254 | 0.909666667 |
| 204 | Pinus sylvestris | 2 | 0.1064 | 0.725666667 |
| 204 | Pinus sylvestris | 5 | 0.2974 | 1.497666667 |
| 204 | Pinus sylvestris | 3 | 0.3234 | 1.648666667 |

Table B.4: Root and shoot dry weights (in grams) collected from plot 204, continued. (Oxberg, Sweden)

# Appendix C

# Data from Studies of Reliability of Visual Estimates

The described data in this section relates to the six studies made in chapter 6.

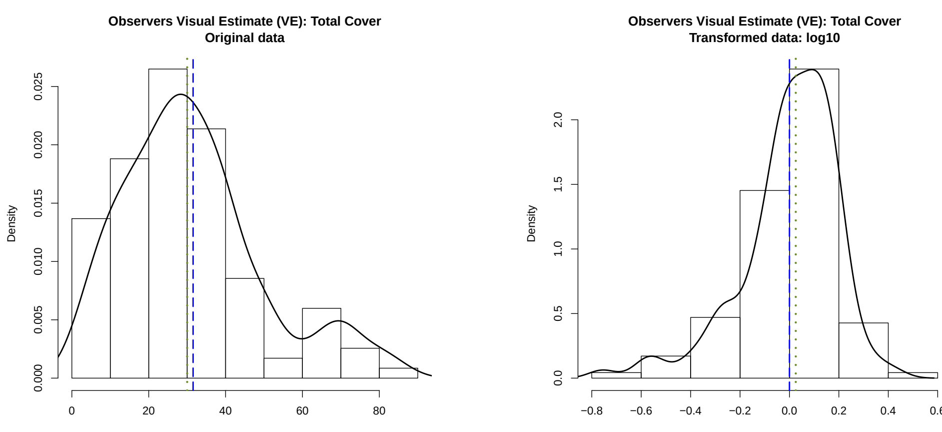
## C.1 Data from the study made in section 6.6

Table C.1 presents the central tendencies median ($Md$) and arithmetic mean ($\overline{x}$) per rater, denoted A to M (i.e. column-wise), and per sample plot, denoted $I1$ to $I9$ (i.e. row-wise), respectively.

| *Obs* | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *J* | *K* | *L* | *M* | *Md* | $\bar{x}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *I1* | 20 | 25 | 10 | 18 | 25 | 10 | 15 | 25 | 15 | 17 | 15 | 15 | 20 | 17 | 17.7 |
| *I2* | 35 | 30 | 25 | 33 | 40 | 30 | 45 | 40 | 35 | 35 | 30 | 40 | 50 | 35 | 36.0 |
| *I3* | 30 | 25 | 20 | 25 | 40 | 25 | 30 | 30 | 25 | 24 | 20 | 25 | 40 | 25 | 27.6 |
| *I4* | 15 | 10 | 10 | 4 | 10 | 4 | 19 | 3 | 7 | 11 | 10 | 2 | 7 | 10 | 8.6 |
| *I5* | 40 | 10 | 15 | 30 | 50 | 75 | 50 | 35 | 40 | 45 | 30 | 35 | 40 | 40 | 38.1 |
| *I6* | 35 | 5 | 15 | 28 | 35 | 40 | 50 | 30 | 37 | 50 | 25 | 30 | 39 | 35 | 32.2 |
| *I7* | 25 | 5 | 10 | 20 | 30 | 20 | 25 | 25 | 30 | 15 | 20 | 25 | 35 | 25 | 21.9 |
| *I8* | 40 | 35 | 35 | 70 | 70 | 85 | 80 | 30 | 60 | 70 | 40 | 40 | 80 | 60 | 56.5 |
| *I9* | 30 | 20 | 15 | 58 | 50 | 70 | 70 | 20 | 65 | 50 | 45 | 30 | 65 | 50 | 45.2 |
| *Md* | 30 | 20 | 15 | 28 | 40 | 30 | 45 | 30 | 35 | 35 | 25 | 30 | 40 | - | - |
| $\bar{x}$ | 30.0 | 18.3 | 17.2 | 31.8 | 38.9 | 39.9 | 42.7 | 26.4 | 34.9 | 35.2 | 26.1 | 26.9 | 41.8 | - | - |

Table C.1: The 13 raters' (A-M) cover estimates (%) from 9 images, I1- I9

The histogram in figure C.1a shows the original visual plant cover estimates given by the raters. A density plot based on the same data as the histogram is superimposed on top of the histogram. Likewise, figure C.1b presents the $\log_{10}$transformed data as a histogram; a density plot is superimposed upon it. The mean is represented with a dashed line, and the median is represented by a dotted line.




a) b)

Figure C.1: a) Histogram and density plots of: a) original data, and b) log10-transformed data

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess whether the residuals were approximately normally distributed (see figure C.2. The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

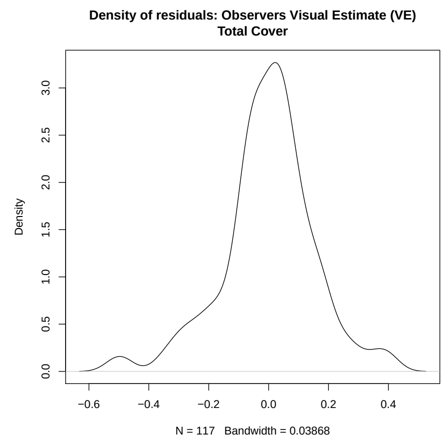



Figure C.2: Density plot of the residuals from the log10-transformed data

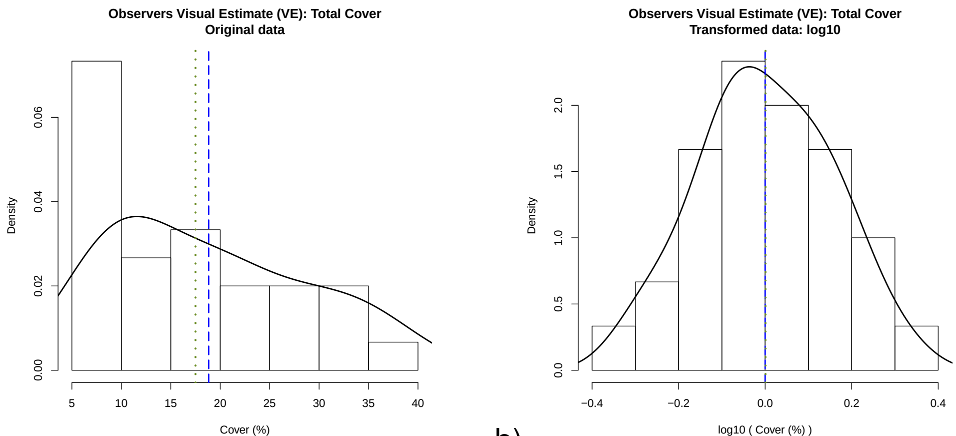
## C.2 Data from the study made in section 6.7

Table C.2 presents all raw data and the central tendencies median ($Md$) and arithmetic mean ($\overline{x}$) per rater, denoted A to C (i.e. column-wise), and per sample plot, denoted $I1$ to $I10$ (i.e. row-wise), respectively.

| *Plot / Observer* | *A* | *B* | *C* | $Md$ | $\overline{x}$ |
|---|---|---|---|---|---|
| $I1$ | 35 | 25 | 10 | 25 | 23.3 |
| $I2$ | 40 | 30 | 25 | 30 | 31.7 |
| $I3$ | 30 | 20 | 20 | 20 | 23.3 |
| $I4$ | 15 | 5 | 5 | 5 | 8.3 |
| $I5$ | 15 | 10 | 10 | 10 | 11.7 |
| $I6$ | 20 | 10 | 15 | 15 | 15.0 |
| $I7$ | 15 | 5 | 10 | 10 | 10.0 |
| $I8$ | 20 | 10 | 10 | 10 | 13.3 |
| $I9$ | 35 | 30 | 35 | 35 | 33.3 |
| $I10$ | 25 | 20 | 10 | 20 | 18.3 |
| $Md$ | 22.5 | 22.5 | 10.0 | | |
| $\overline{x}$ | 25.0 | 16.5 | 15.0 | | |

Table C.2: The three raters (A to C) cover estimates (%) from 10 plots, I1 to I10

The histogram in figure C.3a presents all the original visual plant cover estimates made by the raters. A density plot, based on the same data as the histogram, is superimposed on top of the histogram. Likewise, figure C.3b presents the log10-transformed data as a histogram, with a density plot superimposed upon it. The mean is represented with a dashed line, and the median is represented by a dotted line.




a) b)

Figure C.3: Histogram and density plots of: a) original data, and b) log10-transformed data

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.4). The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

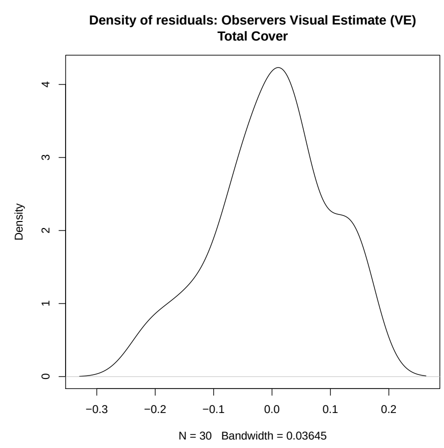
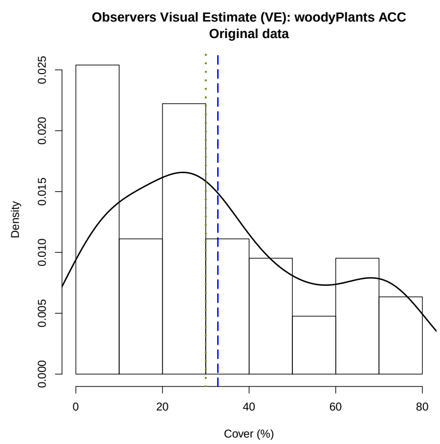



Figure C.4: Density plot of the residuals from the log10-transformed data

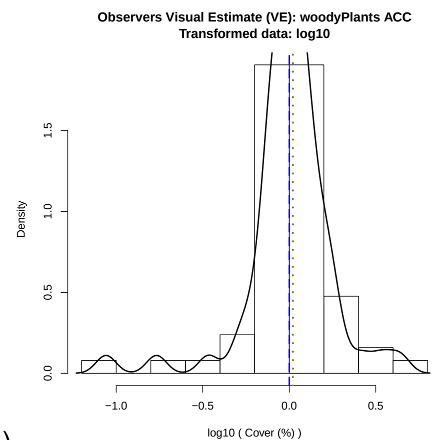
## C.3 Data from the study made in section 6.8

### Visual Estimates of woody plants using ACC method (in section 6.8.2)

The number of observed images in the analysis was 21 out of 35. The rest of the images were not selected because they did not contain any woody plants.




a)




b)

Figure C.5: a) Histogram and density plots of: a) original data, and b) log10-transformed data

The raters' ACC estimate of the woody plant data is presented as a superimposed density plot on a histogram in figures C.5a and C.5b, respectively.

A visual analysis of the histogram (see figure C.5a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis, the data was $\log_{10}$- transformed.

A density plot of the residuals from the $\log_{10}$-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.6). The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

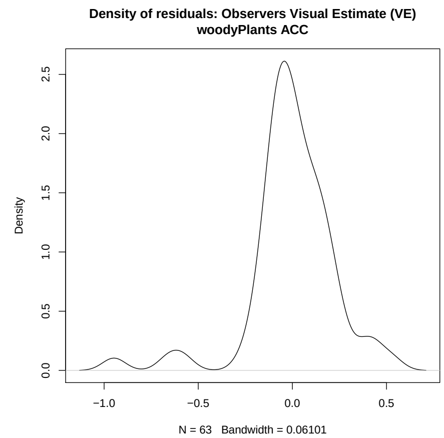
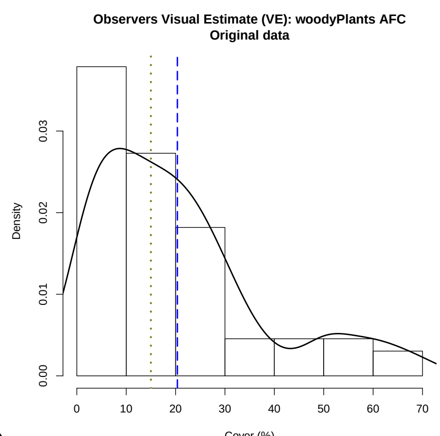


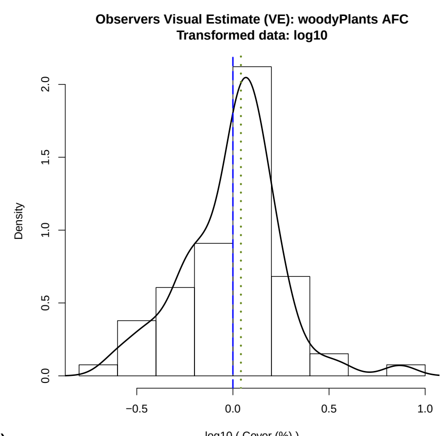

Figure C.6: Density plot of the residuals from the log10-transformed data

**Visual Estimates of woody plants using AFC method (in section** 6.8.2**)**

The number of observed images in the analysis was 22 out of a total of 35. The remaining images were excluded due to the lack woody plants.




a)




b)

Figure C.7: a) Histogram and density plots of a) original data, and b) log10-transformed

The raters' AFC estimate of the woody plant data is presented as a superimposed density plot on a histogram (as shown in in figures C.7a and

C.7b, respectively). A visual analysis of the histogram (figure C.7a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis the data was log10- transformed.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.8).

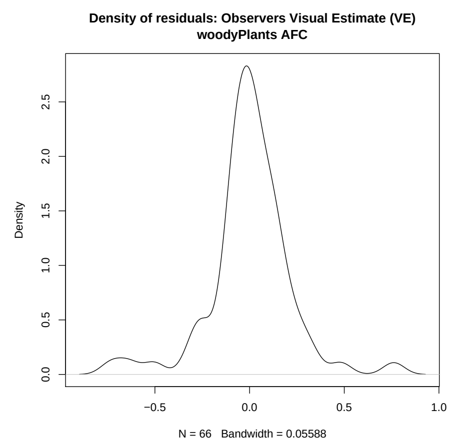
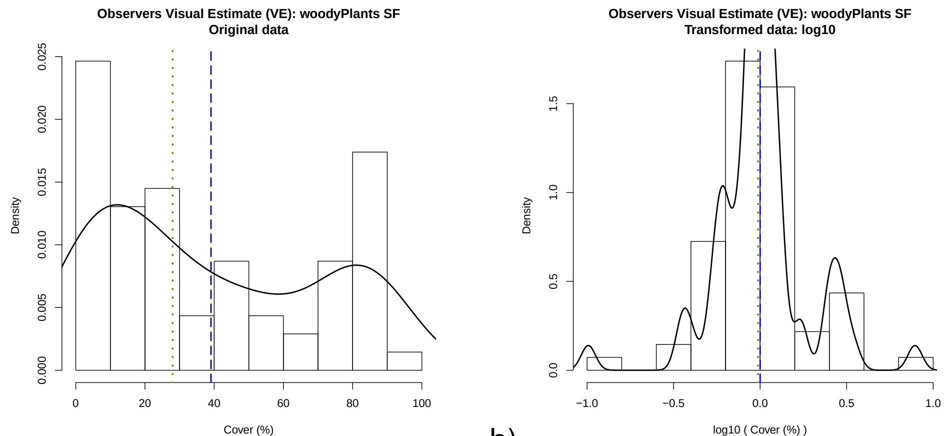



Figure C.8: Density plot of the residuals from the log10-transformed data

**Visual Estimates of woody plants using SF method (in section** 6.8.2**)**

The number of observed images in the analysis was 23 out of 35. The rest of the images were not selected because they did not contain any woody plants.



a) b)

Figure C.9: a) Histogram and density plots of a) original data, and b) log10-transformed data

The raters' SF estimate of the woody plant data is presented as a superimposed density plot on a histogram in figures C.9a and C.9b, respectively. A visual analysis of the histogram (figure6.16a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.10). The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

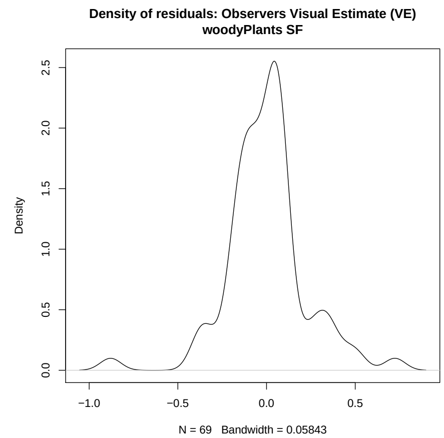
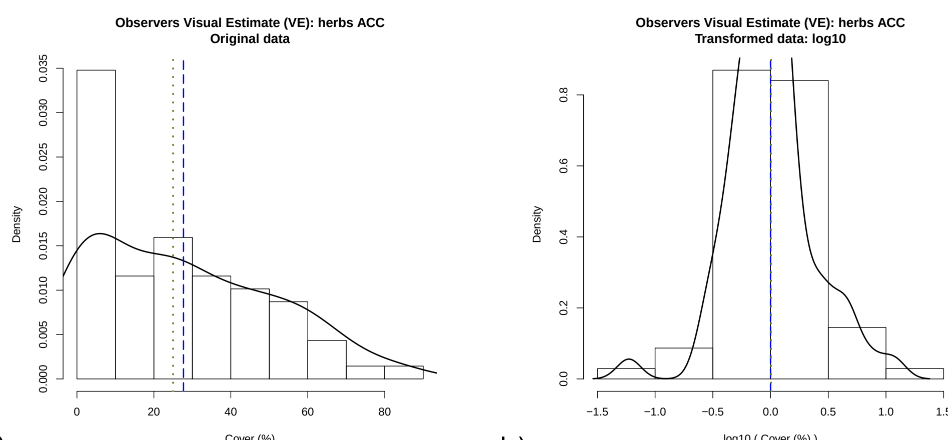



Figure C.10: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Herbs using the ACC Method (in section 6.8.2)

The number of observed images in the analysis was 23 out of 35. In cases where all raters unanimously estimated 0% cover of the target plant, these images were removed.




a) b)

Figure C.11: Histogram and density plots of: a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.12).

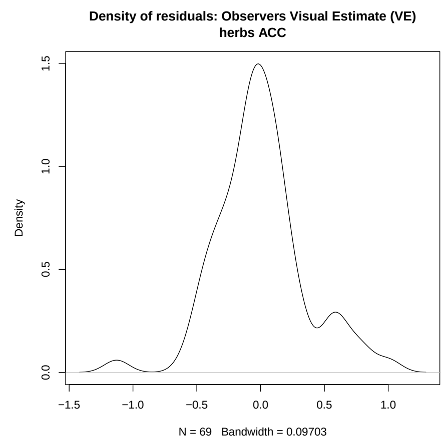
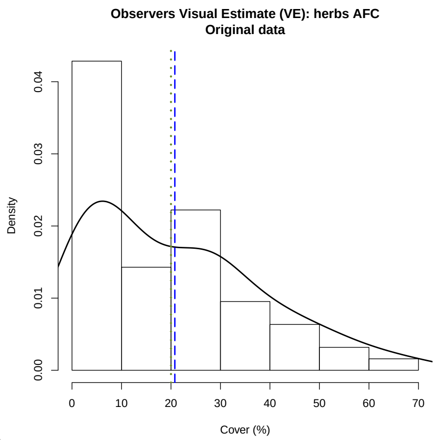


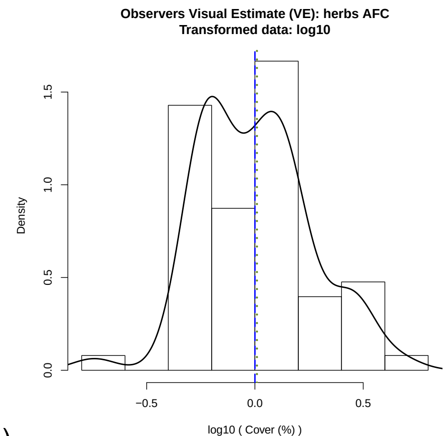

Figure C.12: Density plot of the residuals from the log10-transformed data

**Visual Estimates of Herbs Using the AFC Method (in section 6.8.2)**

The number of observed images in the analysis was 21 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.




a)




b)

Figure C.13: Histogram and density plots of: a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.14).

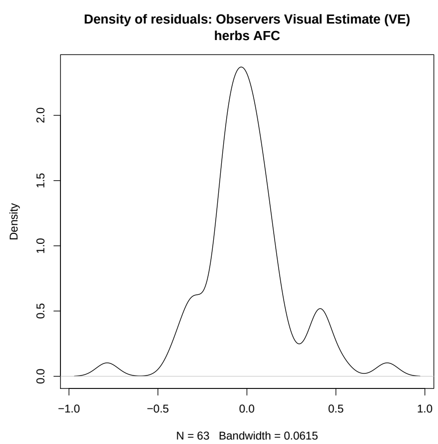
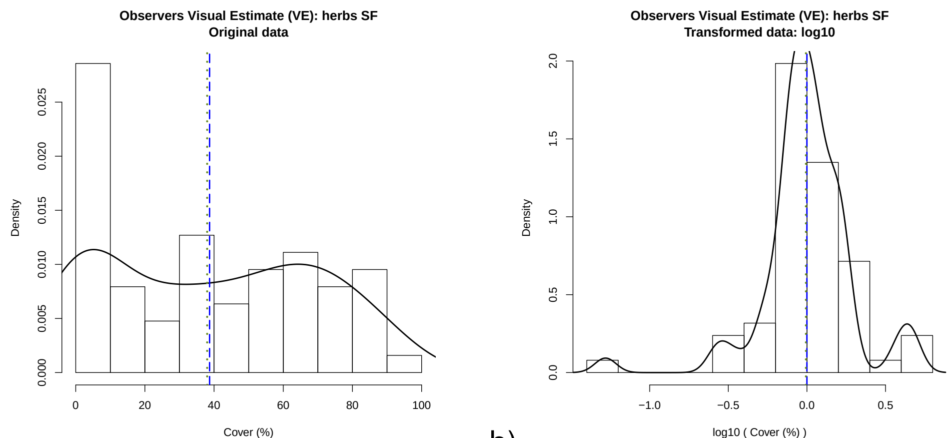



Figure C.14: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Herbs using SF method (in section 6.8.2)

The number of observed images in the analysis was 21 out of 35. The rest of the images were not selected because they did not contain any herbs.




a) b)

Figure C.15: Histogram and density plots of: a) original data, and b) log10-transformed data

The raters' SF estimate of the herbs data is presented as a superimposed density plot on ahistogram in figures C.15a and C.15b, respectively. A visual analysis of the histogram (figure 6.22a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.16). The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

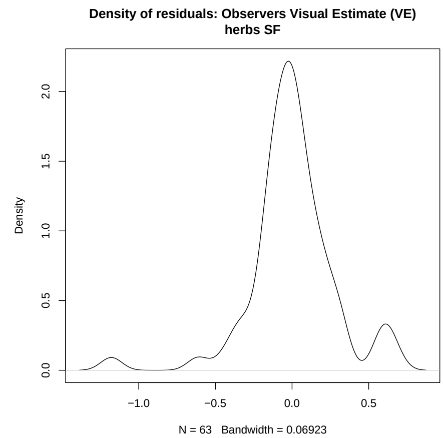



Figure C.16: Density plot of the residuals from the log10-transformed data

## Visual Estimates of Grass Using the ACC Method (in section 6.8.2)

The number of observed images in the analysis was 19 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

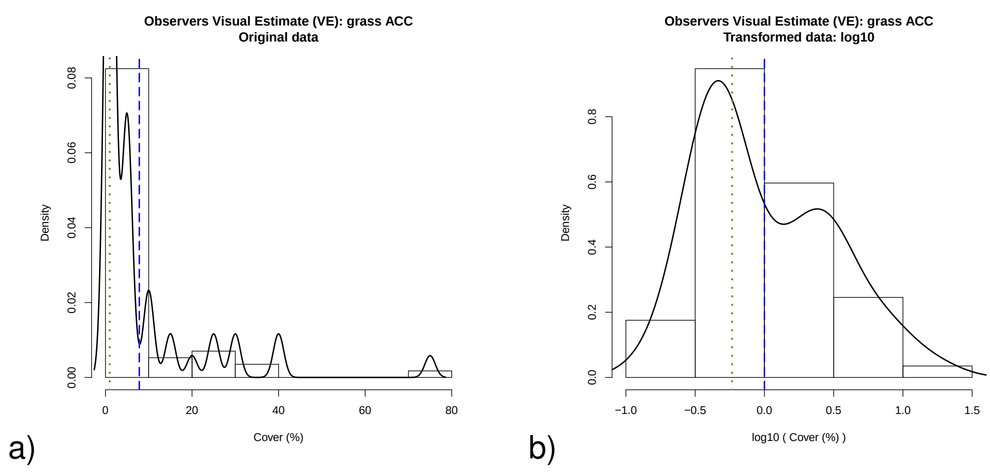



Figure C.17: Histogram and density plots of: a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.18).

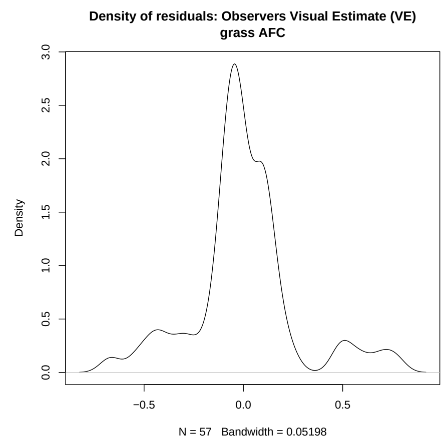



Figure C.18: Density plot of the residuals from the log10-transformed data

**Visual Estimates of Grass using the AFC Method (in section 6.8.2)**

The number of observed images in the analysis was 19 out of 35. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

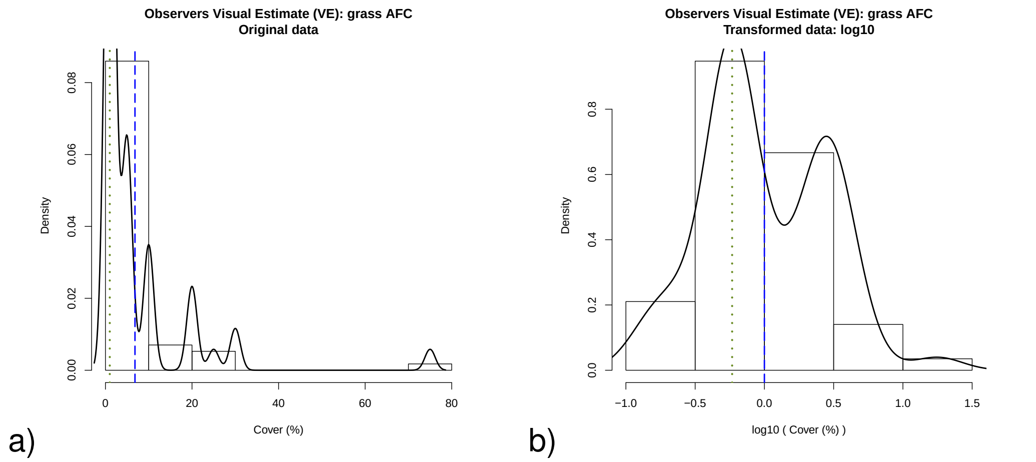



Figure C.19: Histogram and density plots of a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.20).

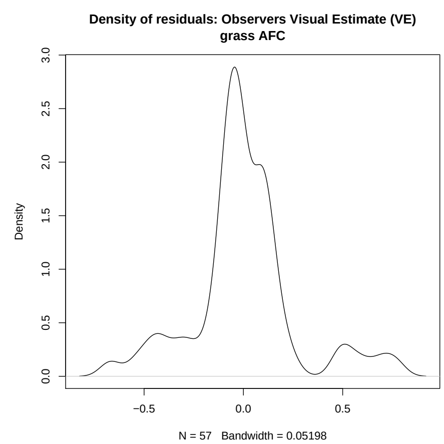



Figure C.20: Density plot of the residuals from the log10-transformed data

## Visual Estimates of Grass using the SF method (in section 6.8.2)

The number of observed images in the analysis was 20 out of 35. The rest of the images were not selected because they did not contain any grass.

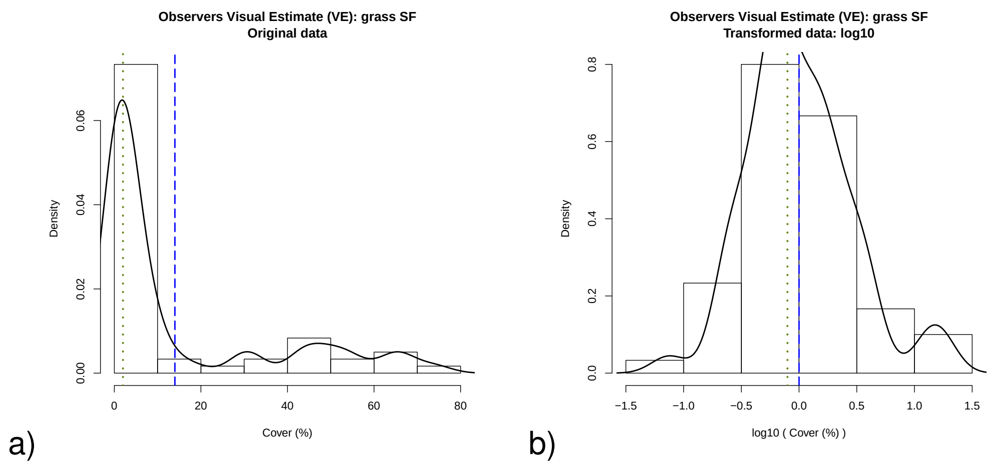



Figure C.21: Histogram and density plots of a) original data, and b) log10-transformed data

The raters' SF estimate of the grass data is presented as a superimposed density plot on a histogram in figures C.21a and C.21b, respectively. A visual analysis of the histogram (figure 6.28a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.22. The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

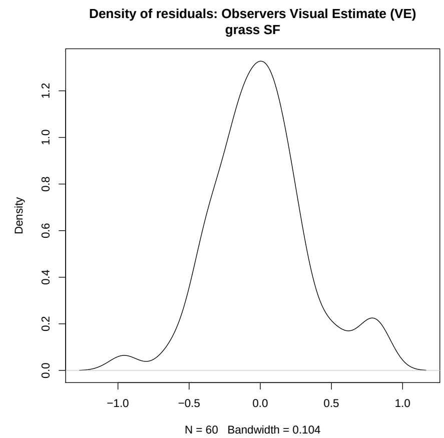
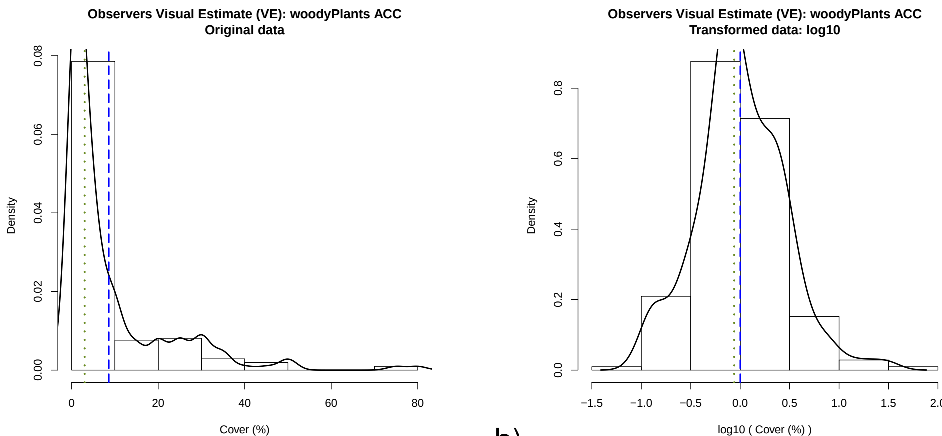



Figure C.22: Density plot of the residuals from the log10-transformed data

## C.4 Data from the study made in section 6.9

### Visual Estimates of Woody Plants Using the ACC Method

The number of observed images in the analysis was 42 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.




a) b)

Figure C.23: Histogram and density plots of: a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figureC.24. The residuals appeared to be approximately normally distributed, thereby justifying the rationale of the choice of ANOVA test.

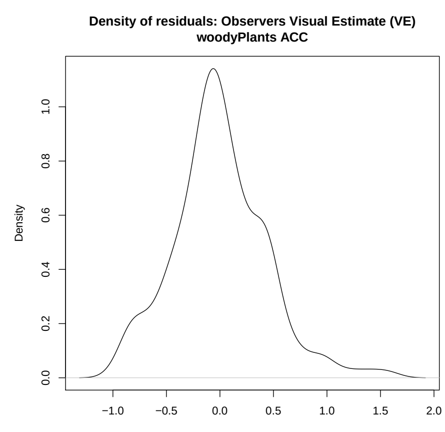
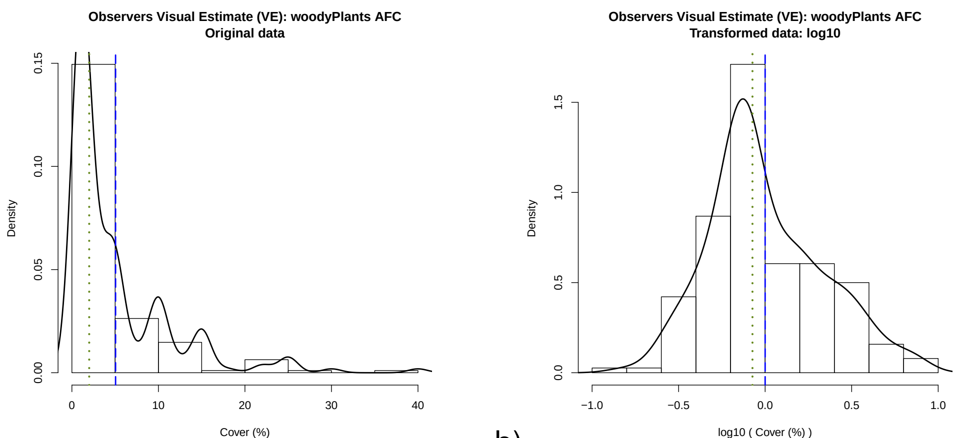



Figure C.24: Density plot of the residuals from the log10-transformed data

## Visual Estimates of Woody Plants Using the AFC Method

The number of observed images in the analysis was 38 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.




a) b)

Figure C.25: Histogram and density plots of: a) original data, and b) log10-transformed

The raters' AFC estimate of the woody plant data is presented as a superimposed density plot on a histogram in figures C.25a and C.25b, respectively. A visual analysis of the histogram (fig. C.25a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis the data was log10-transformed.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.26).

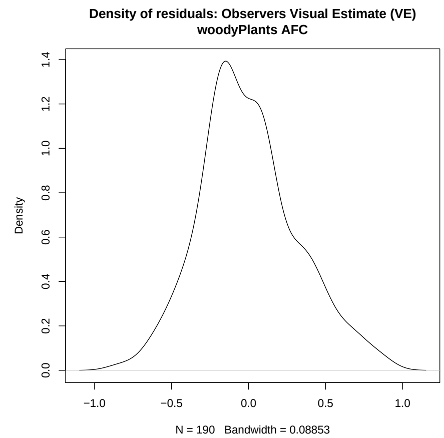



Figure C.26: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Herbs Using the ACC Method

The number of observed images in the analysis was 42 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.

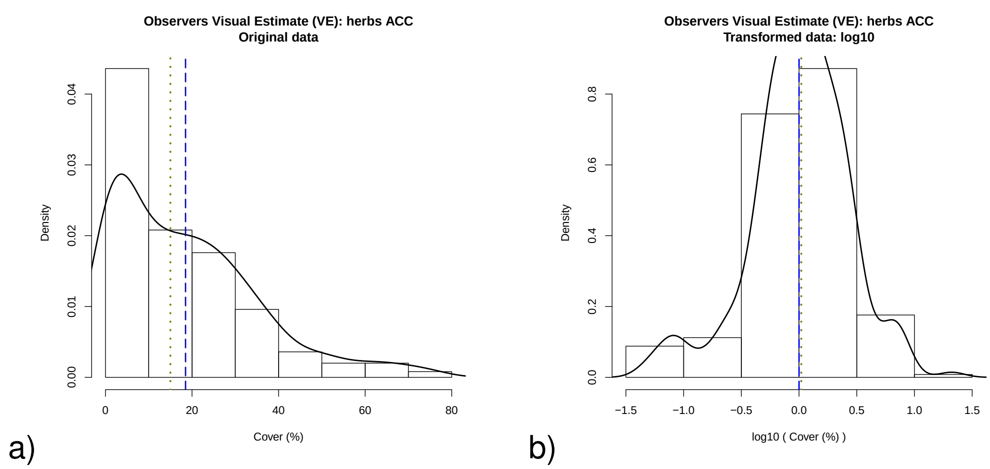



Figure C.27: Histogram and density plots of: a) original data, and b) log10-transformed

The raters' ACC estimate of the herb data is presented as a superimposed density plot on a histogram in figures C.27a and C.27b, respectively. A visual analysis of the histogram (figure C.27a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis the data was log10-transformed.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.28).

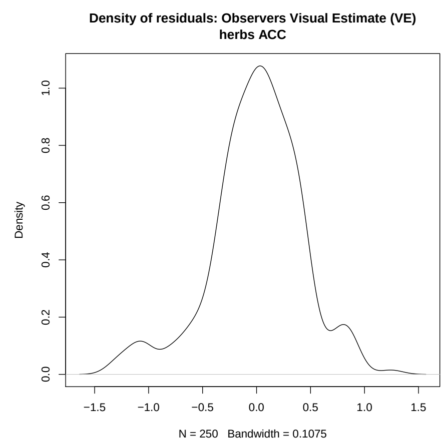
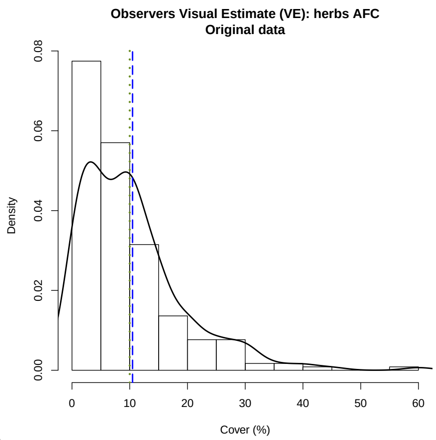


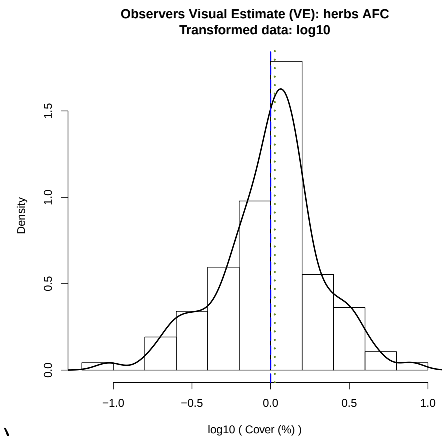

Figure C.28: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Herbs Using the AFC Method

The number of observed images in the analysis was 47 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.



a)



b)

Figure C.29: Histogram and density plots of: a) original data, and b) log10-transformed

The raters' AFC estimate of the herb data is presented as a superimposed density plot on a histogram in figures C.29a and C.29b, respectively. A visual analysis of the histogram (figure C.29a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis the data was log10-transformed.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed, see figure C.30.

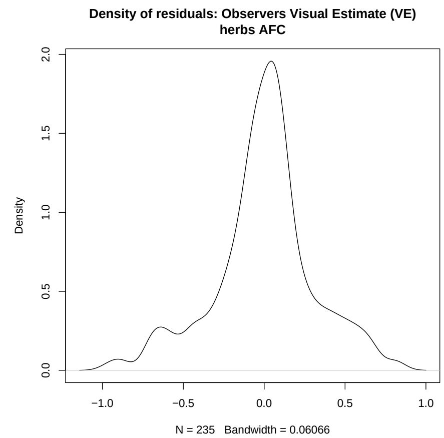
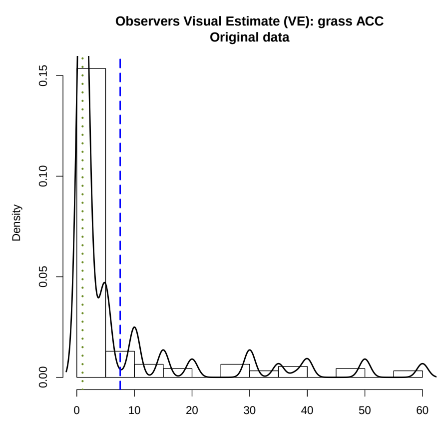


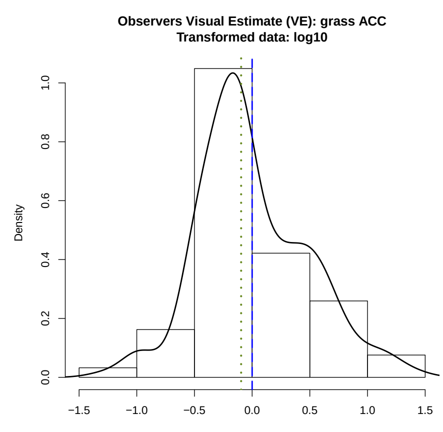

Figure C.30: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Grass Using the ACC Method

The number of observed images in the analysis was 37 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.




a)




b)

Figure C.31: a) Histogram and density plots of a) original data, and b) log10-transformed

As before, the raters' ACC estimate of the grass data is presented as a superimposed density plot on a histogram in figures C.31a and C.31b, respectively. A visual analysis of the histogram (figure C.31a) with the mean plotted as a dashed line and the median plotted as a dotted line indicated an irregular, positively skewed distribution. Therefore, in order to perform a parametric analysis the data was log10-transformed.

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.32).

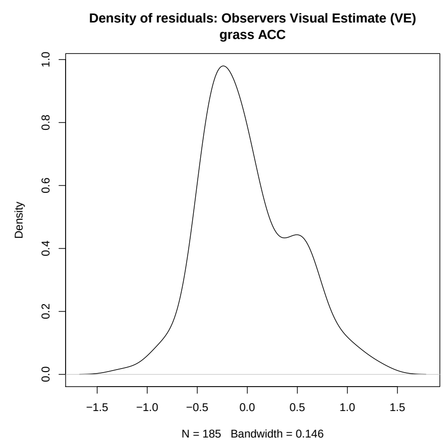
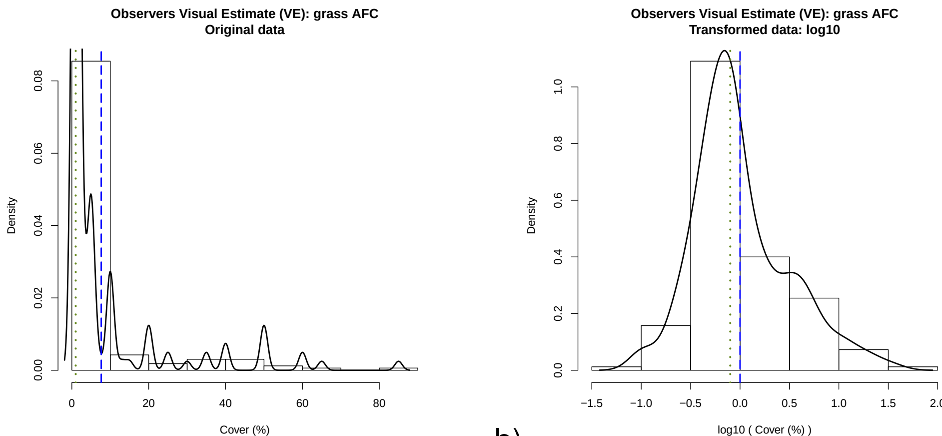



Figure C.32: Density plot of the residuals from the log10-transformed data

### Visual Estimates of Grass using the AFC Method

The number of observed images in the analysis was 33 out of 51. In cases where all raters unanimously estimated an image to contain 0% cover of the target plant, these images were removed.




a) b)

Figure C.33: Histogram and density plots of: a) original data, and b) log10-transformed

A density plot of the residuals from the log10-transformed data was produced to qualitatively assess that the residuals were approximately normally distributed (see figure C.34).

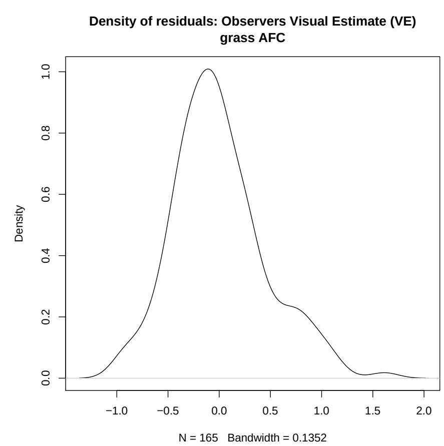



Figure C.34: Density plot of the residuals from the log10-transformed data

# Appendix D

# Botanic Taxonomic Hierarchy

The following table D.1 describes the levels in the taxonomic hierarchy, from the uppermost Kingdom, which is the most general, down to the lowest level Species, which is the most fundamental and precise Mauseth (1998, p.499).

| **Biological Taxonomic Ranks** |
|---|
| Kingdom (regnum) |
| Division (divisio) |
| Class (classis) |
| Order (ordo) |
| Family (familia) |
| Genus (genus) |
| Species (species) |

Table D.1: Botanic taxonomic hierarchy from Mauseth (1998)

# Appendix E

# Experimental Sites: Field Layer Classes and Tree Types

## E.1 General description of experimental locations

This part includes tables that describe the general character of each experimental location, where a location can contain several sites. The tables describe the frequency as a percentage interval for each field layer type or tree type.

| Field layer type | Falun - Grycksbo - Borlänge | Oxberg | Vetlanda |
|---|---|---|---|
| | | | |
| Low herb | 2.1 - 11.0 | 2.1 - 11.0 | 20.1 - 29.0 |
| Tall herb | 10.1 - 18.0 | 10.1 - 18.0 | 10.1 - 18.0 |
| Soil without field layer | 0.6 - 5.0 | < 0.5 | 0.6 - 5.0 |
| Grass | 24.1 - 34.0 | < 8.0 | 34.1 - 47.0 |
| Sedge-horsetail | 1.1 - 4.0 | 1.1 - 4.0 | 4.1 - 7.0 |
| Bilberry (Vaccinium myrtillus) | 24.1 - 36.0 | 12.1 - 24.0 | 12.1 - 24.0 |
| Other shrubs | 5.1 - 16.0 | 5.1 - 16.0 | 5.1 - 16.0 |
| Lichens | 2.1 - 12.0 | > 32.1 | < 2.0 |
| | | | |

Table E.1: Frequencies of field layer (%)

| Tree type | Falun - Grycksbo - Borlänge | Oxberg | Vetlanda |
|---|---|---|---|
| | | | |
| Bare forest land | 6.1 - 9.0 | 3.1 - 6.0 | 6.1 - 9.0 |
| Pine forest | 32.1 - 47.0 | > 62.1 | 17.1 - 32.0 |
| Spruce forest | 24.1 - 38.0 | 10.1 - 24.0 | 38.1 - 52.0 |
| Mixed pine and spruce forest | 16.1 - 21.0 | 11.1 - 16.0 | 16.1 - 21.0 |
| Deciduous forest | 6.1 - 10.0 | 2.1 - 6.0 | 2.1 - 6.0 |
| Mixed deciduous coniferous forests | < 5.0 | < 5.0 | 5.1 - 11.0 |
| | | | |

Table E.2: Frequencies of tree types (%)

# Appendix F

# Database Modelling

## F.1 Database Model: Acquired Data Storage

A relational database model was developed. When implemented, the purpose of the model was to enable the storage of substantial amounts of the acquired data in a structured way. To enable further development and model extensions, the model was normalised down to the third normal form (3NF).

The current database model consists of 10 tables with dependencies and relations, as seen in figure F.1. The tables and their relationships are briefly described here: The table showing sensor types describes all types of sensors, including cameras, laser distance meters, and thermal IR devices. The sensor_types has a 1:M1 relationship with the table of sensors, in which things like sensor manufacturer, and model are shown. Each sensor can have (or produce) several images of different types, or several measurements represented in the tables: image_visible, image_ir, and distance_meter, respectively. Each (camera) sensor can also have much Exif meta data coupled to it. Exif meta data are extensive data about each image, which is the result of having many camera sensors. The data is acquired at data collection points; thus, the table datacollection_ points represents these points. Each data collection point can have many images, or measurements coupled to it. Yet again, these are represented in the tables image_visible, image_ir, and distance_meter. Each data collection point can also have many comments, which are represented in the table comments. When acquiring data at the collection points, the weather conditions are often of interest. The table present_weather_type contains World Meteorological Organization (WMO) standard weather types, to be used as a record in the table weather_observation, as well as additional observed weather attributes.

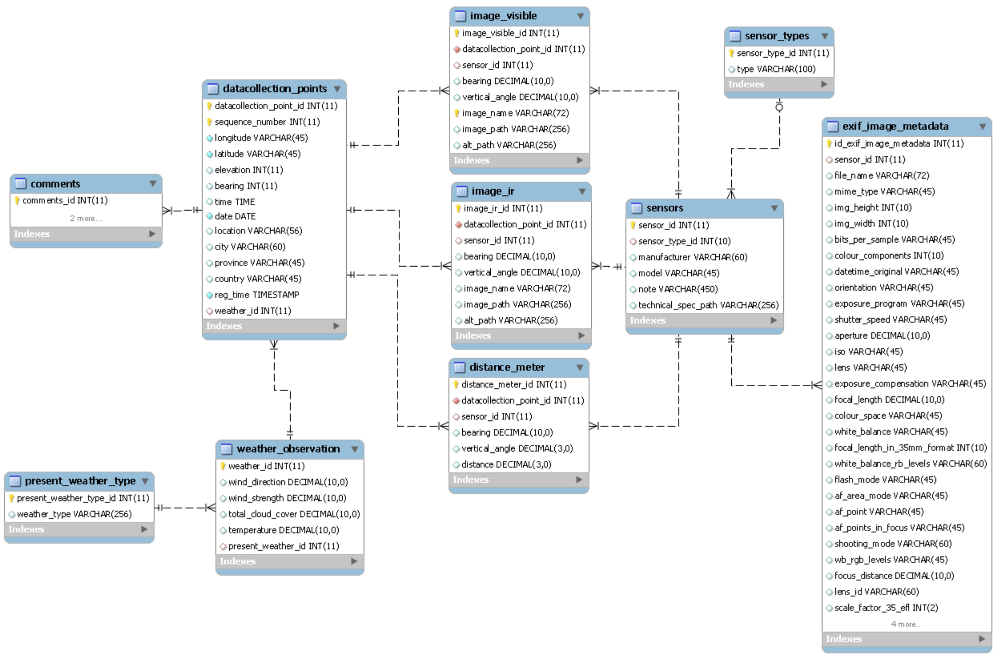



Figure F.1: Database model

# Appendix G

# Assistance when Estimating Plant Cover

Another type of template that complements figure 6.10 and enables the raters to make better and more synchronous estimates of cover is shown in figure G.1. The templete outlines three different percentages of cover: 5, 10 and 25%, respectively. The three sketches on the first row are all 5%, the three sketches on the second row are all 10%, and so on.

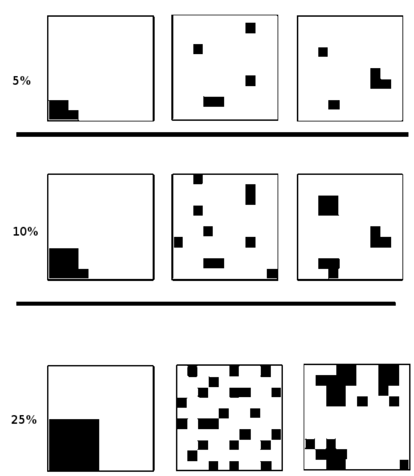



Figure G.1: Aiding template for the estimate of plant cover

# Appendix H

# Differences in Counting Plants Clusters

The graphical plots in this appendix (see figure H.1) all describe the difference between two sessions in which the number of plants clusters were estimated. Five raters (denoted A to E) participated. Each rater counted plant clusters from 51 images twice in separate sessions. The results and conclusion are described in section 6.9.

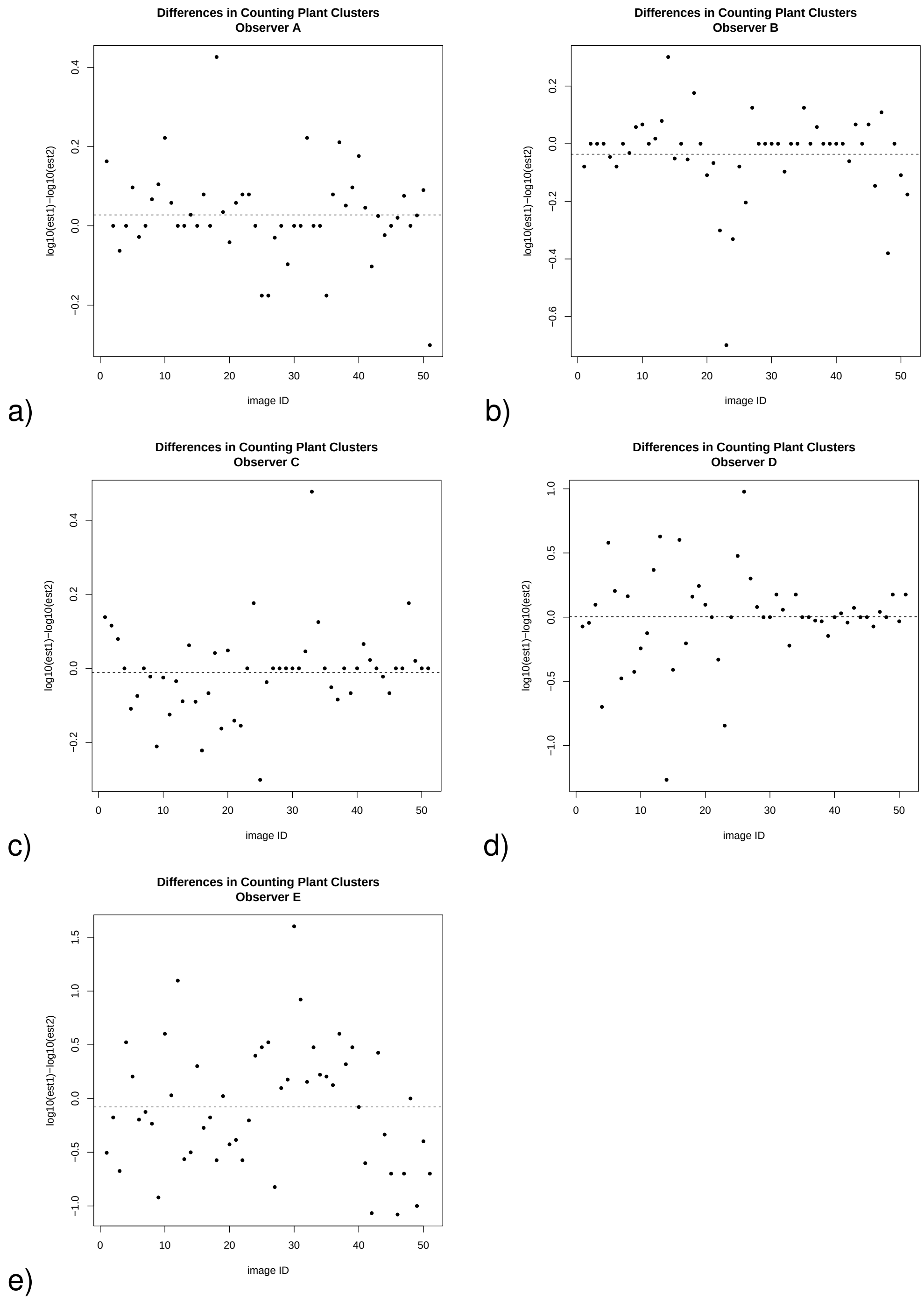



Figure H.1: Differences in counting plant clusters between rater A to E

# Appendix I

# The Engineering Design Process & The Scientific Method

In this appendix, the traditional scientific method and the engineering design process are briefly described (see table I.1). Steps 1 to 6 do not necessarily flow in a sequential order, but are often iterative. The results are presented as step 7.

| Step | **The Engineering Design Process** | **The Scientific Method** |
|---|---|---|
| 1 | Define a problem or need | State a question or problem |
| 2 | Gather background information | Gather background information |
| 3 | Establish design statement or criteria | Formulate hypothesis; identify variables |
| 4 | Prepare preliminary designs | Design experiment, establish procedure(s) |
| 5 | Build and test a prototype(s) | Test hypothesis by doing an experiment |
| 6 | Verify, test & redesign as necessary | Analyse results & draw conclusions |
| 7 | Present results | Present results |

Table I.1: The engineering design process and the scientific method

# Appendix J

# Flow-chart for estimating the negative binomial exponent $k$

The information provided in this appendix are based on the four suggested methods, in Krebs (1999), for estimating the negative binomial exponent $k$. An estimate of the negative binomial exponent $k$ is needed when one wants to describe (plant) populations that show aggregated (or clumped) patterns (see section 4.2 for more details about aggregated patterns and the negative binomial distribution).

## J.1 Methods

The methods $\sharp$1 to $\sharp$4, below, refers to the flow-chart in figure J.1.

**$\sharp$1 Calculate approximate $k$ (Large samples: Quadrats$\geq 20$)**

Use equation 4.7 in section 4.2.

**$\sharp$2 *More than 1/3 of the quadrats are empty* (Small samples: Quadrats$< 20$)**

Calculate an estimate of $k$ by solving equation J.1 iteratively (i.e. by trial and error):

$$log_e\left(\frac{N}{n_0}\right) = \hat{k} * log_e\left(1 + \frac{\overline{x}}{\hat{k}}\right) \tag{J.1}$$

where $N$= total number of quadrats counted; $n_0$ = number of quadrats containing zero individuals; $\overline{x}$= observed mean; and$\hat{k}$= estimate of the negative binomial exponent

Begin with the approximate value of k calculated above and raise or lower it to make the two sides of this equation balance.

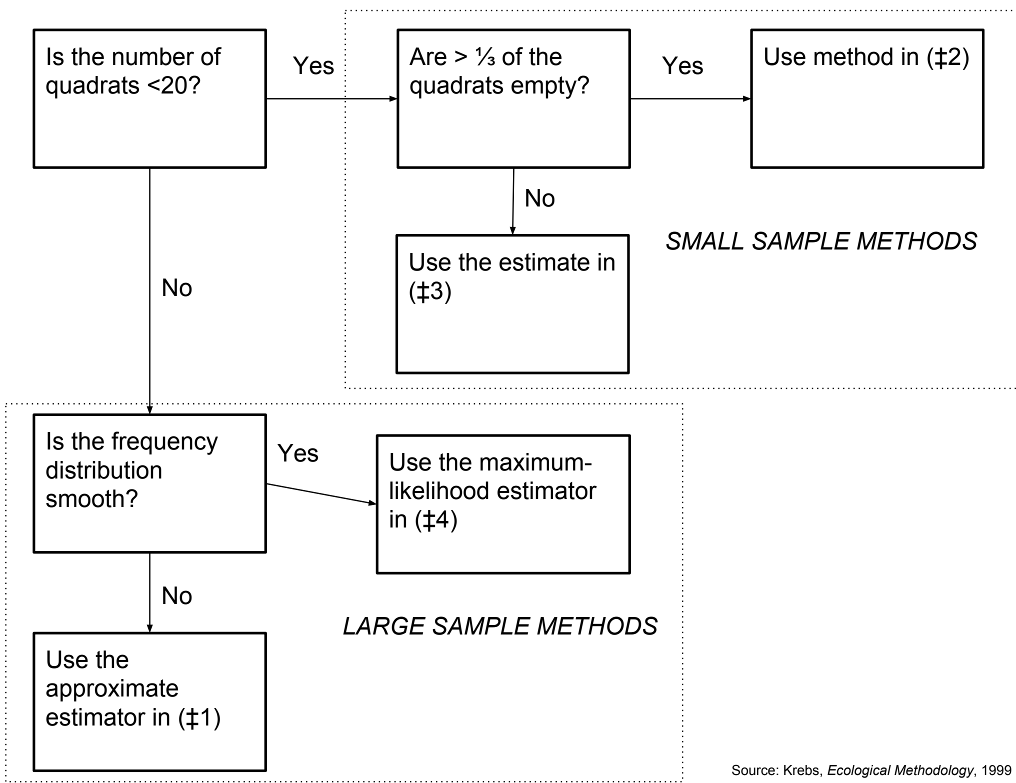



Figure J.1: Estimating the negative binomial exponent $k$, with reference to method ‡1 to ‡4 in section J.1.

**♮3 *Less than 1/3 of the quadrats are empty* (Small samples: Quadrats$< 20$)**

$$\hat{k} = \frac{\bar{x}^2 - \left(s^2/n\right)}{s^2 - \bar{x}} \tag{J.2}$$

**♮4 Smooth frequency distribution with no extremely large counts (Large samples: Quadrat $\geq 20$)**

Calculate a maximum likelihood estimate for $k$ by solving equation J.3, by trial and error:

$$(N)\, log_e \left(1 + \frac{\bar{x}}{\hat{k}}\right) = \sum_{i=0}^{\infty} \left(\frac{A_x}{\hat{k} + x}\right) \tag{J.3}$$

where $N$ = total number of quadrats counted; $\bar{x}$ = observed mean; $\hat{k}$ = estimated negative-binomial exponent

$A_x = \sum_{j=x+1}^{\infty} \left(f_j\right) = f_{x+1} + f_{x+2} + f_{x+3} + ...$, and $i$ = a counter $(0, 1, 2, 3......)$

# Appendix K

# Internet-based Sources

The following Internet sources, most often the start of research databases, were used during the reviews. Journal and conference papers of interest were downloaded, or ordered as paper copies on behalf of the libraries at Edinburgh Napier University, Scotland, or Dalarna University, Sweden. In addition, books and technical reports were used for the review.

| Name of resource | Address |
|---|---|
| IEEE Xplore | http://ieeexplore.ieee.org |
| ACM | http://www.acm.org |
| Springer Link | http://link.springer.com |
| Wiley Online Library | http://onlinelibrary.wiley.com |
| JSTOR | http://www.jstor.org |
| Science Direct (Elsevier Journals) | http://www.sciencedirect.com |
| Emerald Journals | http://www.emeraldinsight.com |
| Web of Science | http://apps.webofknowledge.com |
| SAGE Journals | http://online.sagepub.com |
| Oxford Reference Online | http://www.oxfordreference.com |

Table K.1: Internet-based sources in the literature review